



The Alberta Journal of Educational Research

Volume XXI, No. 1

March, 1975

CONTENTS

The judgment of simple deductive arguments by pre-service elementary school teachers	1
<i>L. C. Jansson</i>	
Validity of Expectation—A Developmental Function	11
<i>M. M. Clifford</i>	
The Interview Reassessed	18
<i>L. Foster and M. Nixon</i>	
Does the Dolch Do?	23
<i>V. Froese</i>	
Studying Problem Solving Behavior in Young Children—Some Methodological Considerations	28
<i>D. Nelson</i>	
Controlled vs. Natural Setting—Some Implications for Behavioral Analysis and Change in Classroom Situations	39
<i>J. Martin</i>	
Day Care Teacher Practices and Beliefs	46
<i>S. Verma and D. L. Peters</i>	
Learning the Language of Reading—An Exploratory Study	56
<i>A. D. Forester</i>	
The Works of Jean Piaget Published in English	63
<i>B. M. Blackall</i>	

FACULTY OF EDUCATION
The University of Alberta

L . C . J A N S S O N

University of Manitoba

The Judgment of Simple Deductive Arguments by Pre-Service Elementary School Teachers

The purpose was the assessment of pre-service elementary school teachers' abilities to judge selected deductive arguments. 140 Ss were given tests requiring validity judgments of five varieties of argument in isomorphic class and conditional forms. The judgment of arguments of the class type were found to be significantly easier than the judgment of arguments of the conditional type. A significant difference was also found among abilities to judge the validity of the five varieties (principles) of argument. Furthermore, it was possible to rank these five principles in order of difficulty, and the ranking was constant over the two types of reasoning. A significant interaction of type and principle was found. The percentage of Ss mastering each of the five principles was also determined. (Dr. Jansson is Associate Professor in the Department of Curriculum, Faculty of Education, The University of Manitoba.)

Over the past decade a number of researchers, *e.g.*, Hill (1961), Howell (1965), O'Brien and Shapiro (1968), O'Brien, Shapiro, and Reali (1971), have examined the developmental aspects of children's abilities to handle deductive arguments, valid and invalid, of simple sentential and class types. These studies have been concerned with logical skills *per se*.

In addition to the studies cited above, there is a set of investigations spawned by the work of the Cornell Critical Thinking Project. Some of these (Gardiner, 1966; Paulus, 1967; Roberge, 1970) have, like the above studies, looked at developmental patterns in logical skills development, while others (Ennis, 1959 and 1962; Jansson, 1970) have focussed on the broader notion of critical thinking ability, an important component of which is the ability to assess the validity of logical arguments.

Hill (1961) suggested that children of ages six, seven, and eight "are able to deal very effectively with verbal premises that call for hypothetical reasoning . . .," to which O'Brien and Shapiro added some important qualifications based upon their more open response form. The latter study

used the same item types (sentential logic, classical syllogism, and logic of quantification) as Hill, but included a "not enough clues" distractor, and concluded that the abilities found by Hill "cannot be taken for granted in children of this age." Neither of these studies involved explicit instruction in the logical abilities. Roberge and Paulus (1971) also used an assessment form in their investigation of children's abilities in class and conditional reasoning. Overall in grades 4, 6, 8, and 10 they found a significant ($p < .05$) difference in children's abilities to handle the two types of arguments. However, significant differences did not appear at each grade level tested, and in fact, the means were reversed at grade 8 in favor of the conditional, rather than class reasoning. The complete analysis showed a strong interaction between type of reasoning and a three-level content dimension (concrete-familiar, abstract, and symbolic). As these researchers state, the "results suggest that it is psychologically and pedagogically feasible to introduce brief units on class and conditional reasoning as early as the fourth grade." They also indicate possible sources of difficulty (e.g., content) toward which teaching strategies could be directed. "Such information should be valuable to mathematics teachers, curriculum planners, and psychologists alike (p. 199)."

The Present Study

The focus in the present study is on the determination of an individual's ability to assess a logical argument, rather than on his ability to deduce. It is pointed out by Paulus (1967) that these abilities are different (deducing is more difficult) in certain important respects, but that the difference seems to decrease as the students move from grade 5 to grade 11, the range covered in his study. Again, there are important interactions with other variables such as content and form, but this is not of immediate concern here.

All of the studies cited here, while they have many points of disagreement, appear to suggest that some instruction in logical reasoning may be appropriate by at least grade 4, if not before. In fact, such materials is finding its way into elementary mathematics textbook series. This has important implications for our teacher training programs.

The above discussion and the cited studies themselves give some insight into the status of children's logical reasoning abilities. The present study was formulated as a status study of pre-service elementary teachers' abilities to judge logical arguments employing five specified principles of reasoning in both class and conditional forms. Clearly teachers must possess these skills if they are going to provide explicit (or even implicit) instruction in them.

More specifically, answers were sought to the following questions:

1. Do the teachers' abilities to judge deductive arguments vary according to the type of reasoning?
2. Do the teachers' abilities to judge deductive arguments vary according to the principle of reasoning?
3. Do the teachers' abilities to judge deductive arguments in class and conditional form vary according to the principle?
4. What proportion of the teachers have mastered each of the five principles of logical reasoning?

In each of the above questions "teachers" refers to pre-service elementary school teachers taking a regular elementary education mathematics methods course. Each subject has had at least one course in mathematics which may have included some material on formal logic. The answers to the questions should prove valuable in structuring a unit on logic for such students and in providing instructional hypotheses relative to their learning of logical skills.

Method

Tests

The tests used in this investigation were:

1. *The Roberge Class Reasoning Test*, Form AN
2. *The Roberge Conditional Reasoning Test*, Form BN

Both were developed by Roberge and used with permission. These instruments have evolved from earlier ones used by Roberge and Paulus and stem from the conceptual and empirical work of the Cornell Critical Thinking Project.

The test items evaluated the subjects' abilities on five principles of deductive reasoning. Item format is as follows:

Suppose you know that

premise

premise

Then would this be true?

Conclusion

Three responses were possible and were defined as follows for the examinees:

- | | |
|----------|---|
| 1. YES | It must be true. |
| 2. NO | It can't be true. |
| 3. MAYBE | It may be true or it may not be true. You weren't told enough to be <i>certain</i> whether it is "YES" or "NO." |

Principles

The five principles of reasoning of each type are exhibited in Tables 1 and 2. More background information on these and related tests may be found in Paulus (1967), Roberge (1970), and Roberge and Paulus (1971).

Procedures

The sample of 140 subjects was chosen randomly from approximately 200 junior and senior undergraduates enrolled in an elementary mathematics methods course during the fall of 1972. The tests were administered during the first two weeks of the term, one week apart. Half of the group received the Conditional Reasoning test first, the other half took the Class Reasoning test first.

In order to answer the questions posed, a 2x2x5 analysis of variance (fixed factors) with repeated measures on the last two factors was performed (Winer, 1962, 319-337). The first factor was that of test sequence. The second and third factors, the ones of critical interest, were, respectively, types of reasoning, and principles. The Newman-Keuls method was used for *post hoc* comparisons of means. Further analysis of the data indicated the proportion mastering each of the principles.

TABLE 1
BASIC PRINCIPLES OF CLASS REASONING

Principle	Validity	Symbolic Form	Concrete Form
1	YES	All A's are B's x is an A ∴ x is a B	All of the dogs that are brown are named Rover. The dog is brown. ∴ The dog's name is Rover.
2	NO	All A's are B's x is a B ∴ x is an A	All of the paintings that belong to Don are paintings of horses. This is a painting of a horse. ∴ This painting belongs to Don.
3	NO	All A's are B's x is not an A ∴ x is not a B	All of the cars in the garage that are Mr. Smith's are black. The car in the garage is not Mr. Smith's. ∴ The car is not black.
4	YES	All A's are B's x is not a B ∴ x is not an A	All of the cats that are not named Tabby are also not white. The cat is white. ∴ The cat's name is Tabby.
5	YES	All A's are B's All B's are C's ∴ All A's are C's	All of the coats in the closet that belong to Donna are blue. All of the coats that are blue have white buttons. ∴ All of the coats in the closet that belong to Donna have white buttons.

Results and discussion

The dependent variable in the analysis of variance was subjects' scores on the five principle subtests. Since each subtest contained eight items, possible scores ranged from 0 to 8. The analysis of variance is presented in Table 3. Tables 4, 5 and 6 present means and differences between means while Table 7 is a summary of the proportions of subjects mastering each principle.

The F ratios for the factor test sequence (Table 3) and its interactions with the factors types of reasoning and principles were non-significant ($p > .05$).

The significant ($p < .05$) F ratio for the factor types of reasoning indicated a significantly higher mean in class reasoning than in conditional reasoning when considered together with the total mean scores for each type (Table 4).

The significant ($p < .001$) F ratio for the factor principles of reasoning indicates that abilities varied significantly among the five principles of reasoning. The means given in the rows of Table 4 entitled "combined" refer to means of all individuals over each principle, weighted proportionally to standard deviation, *i.e.*, scores are added and divided by two. The

TABLE 2
BASIC PRINCIPALS OF CONDITIONAL REASONING

Principle	Validity	Symbolic Form	Concrete Form
1	YES	if p, then q $\frac{p}{q}$	If the dog is brown, then his name is Rover. The dog is brown. The dog's name is Rover.
2	NO	if p, then q $\frac{q}{p}$	If the painting belongs to Don, then it is a painting of a horse. This is a painting of a horse. This painting belongs to Don.
3	NO	if p, then q $\frac{\text{not } p}{\text{not } q}$	If the car in the garage is Mr. Smith's, then it is black. The car in the garage is not Mr. Smith's. The car is not black.
4	YES	if p, then q $\frac{\text{not } q}{\text{not } p}$	If the cat's name is not Tabby, then she is not white. The cat is white. The cat's name is Tabby.
5	YES	if p, then q if q, then r $\frac{\text{if p, then r}}{\text{if p, then r}}$	If the coat in the closet belongs to Donna, then it is blue. If it is blue, then it has white buttons. If the coat in the closet belongs to Donna, then it has white buttons.

TABLE 3
ANALYSIS OF VARIANCE FOR TEST SEQUENCE, TYPES OF REASONING
AND PRINCIPLES OF REASONING

Source	df	MS	F
<u>Between Subjects</u>			
A (Test Sequence)	1	0.93	0.10
Subjects Within Groups	138	9.49	
<u>Within Subjects</u>			
J (Types of Reasoning)	1	10.29	4.97*
A x J	1	7.73	3.74
J x Subjects Within Groups	138	2.07	
K (Principles of Reasoning)	4	813.23	172.341**
A x K	4	2.34	0.496
K x Subjects Within Groups	552	4.72	
J x K	4	13.77	9.170**
A x J x K	4	8.22	5.476**
J x K x Subjects Within Groups	552	1.50	

* $p < .05$
** $p < .001$

TABLE 4
CROSS TABULATION OF MEANS FOR ANALYSIS OF VARIANCE

Test Sequence	Type of Reasoning	Principle					
		1	2	3	4	5	All
Conditional - Class	Conditional	6.67	2.94	3.59	4.60	6.14	23.94
	Class	7.35	3.40	3.94	4.59	6.26	25.54
	Combined	7.01	3.17	3.76	4.59	6.20	24.74
Class - Conditional	Conditional	7.36	2.30	3.47	4.76	6.54	24.43
	Class	7.07	3.50	3.70	4.47	5.80	24.54
	Combined	7.21	2.90	3.59	4.61	6.17	24.49
Both	Conditional	7.01	2.62	3.53	4.68	6.34	24.19
	Class	7.21	3.45	3.82	4.53	6.03	25.04
	Combined	7.11	3.04	3.68	4.60	6.19	24.62

analysis of pairwise comparisons (Table 5) indicated that all are significant and that thus the principles can be ranked in order of difficulty: 2,3,4, 5,1. This ranking applies both to the combined scores and within each type. Although this order is somewhat different than that found by Roberge (1970) for tenth graders: 3,2,4,5,1, the reason for the shift is not clear. Differences between means for the remaining types X principles simple main affects comparisons are presented in Table 6. Although few unexpected patterns emerge clearly, three characteristics of the matrix are noteworthy. First, an examination of the main diagonal reveals significant differences between class and conditional reasoning types for both invalid principles, affirming the consequent and denying the antecedent. Both principles appear more difficult to judge in the conditional reasoning form. Secondly, the familiar transitivity property appeared easier in the conditional form. Finally, the judgment of a *modus ponens* form of argument of the class type was significantly easier than all other arguments except *modus ponens* itself in the conditional form.

A further analysis was made to determine the proportion of subjects who had mastered each of the five principles. These results are presented in Table 7. Each principle subtest contained eight items and mastery was defined as 6/8. Using cumulative binomial probabilities, it can be determined that the probability of showing mastery of a single principle of one type by chance alone is approximately .02. The fourth column of Table 7, headed "Both" gives the proportion who mastered the principle on both class and conditional forms.

It is instructive to compare the results obtained in this study with those of previous studies. In a recent summary of research related to children's comprehension of deductive reasoning schemes Roberge (1972) presented a table showing percentages of students who had mastered each of the prin-

TABLE 5
DIFFERENCES FOR NEWMAN-KEULS MAIN EFFECTS
AND SIMPLE MAIN EFFECTS COMPARISONS

Pairs of Principles	Types		
	Conditional	Class	Combined
1 & 2	4.39	3.76	4.08
1 & 3	3.49	3.39	3.44
1 & 4	2.69	2.34	2.51
1 & 5	1.19	.67	.93
2 & 3	.37	.91	.64
2 & 4	1.08	2.08	1.57
2 & 5	2.58	3.72	3.15
3 & 4	.71	1.15	.93
3 & 5	2.21	2.81	2.51
4 & 5	1.50	1.06	1.58

All significant at $p < .05$

TABLE 6
DIFFERENCES FOR NEWMAN-KEULS MAIN EFFECTS COMPARISON
PRINCIPLES AND TYPES

Class						
Conditional	Ordered Means	K ₂	K ₃	K ₄	K ₅	K ₁
		3.45	3.82	4.53	6.03	7.21
K ₂	2.62	-.83*	-1.20*	-1.91*	-3.41*	-4.59*
K ₃	3.53	.08	-.29*	-1.00*	-2.50*	-3.69*
K ₄	4.68	1.23*	.86*	.15	-1.35*	-2.54*
K ₅	6.34	2.89*	2.52*	1.81*	.31*	-.87*
K ₁	7.01	3.56*	3.19*	2.49*	.99*	-.20

$p < .05$

Note: Table shows differences in principle means (Conditional-Class);
negative indicates class has larger mean, hence easier.

TABLE 7
PROPORTIONS OF SUBJECTS MASTERING EACH PRINCIPLE

Principle	Proportion Mastering		
	Conditional	Class	Combined
1	.84	.91	.76
2	.11	.19	.08
3	.21	.33	.17
4	.31	.27	.11
5	.78	.65	.57
All	.00	.00	.00

TABLE 8
PERCENTAGES OF STUDENTS WHO HAD MASTERED EACH OF FIVE
PRINCIPLES OF DEDUCTIVE REASONING IN CONDITIONAL FORM
(Roberge, 1972)

Grade Level	Principles					n/N	p	Researcher
	1	2	3	4	5			
4	53	2	2	35	28	8/12	.02	Roberge (1970)
5	51	2	3	30	25	5/6	.02	Ennis & Paulus (1965)
4-5	59	0	2	16		5/6	.02	Gardiner (1965)
6	54	0	0	23	28	8/12	.02	Roberge
7	56	3	6	41	45	5/6	.02	Ennis & Paulus
6-7	79	0	2	23		5/6	.02	Gardiner
8	95	0	0	74	81	8/12	.02	Roberge
9	66	4	5	35	40	5/6	.02	Ennis & Paulus
8-9	89	4	4	48		5/6	.02	Gardiner
10	100	19	5	65	82	8/12	.02	Roberge
11	62	3	12	35	58	5/6	.02	Ennis & Paulus
10-11	99	30	20	47		5/6	.02	Gardiner
-----	---	---	---	---	---	---	---	-----
Pre-service Elem. Tchg. Jrs. & Srs.	84	11	21	31	78	6/8	.02	Jansson

n=number of items required for mastery of a principle
N= total number of items for a principle
p= probability of correctly answering at least n items by guessing alone.

ciples of inference. Actually Roberge in this summary lists published results for deductive reasoning only in the conditional form. It is clear both in Roberge's earlier work (1970) and in the present study that the measured ability to handle a particular principle is different if the argument is presented in the class form. In Table 8 part of Roberge's (1972, p. 199) table is reproduced, showing the percentages of students who had mastered a given principle at each grade level.

Table 9 presents similar mastery data for deductive reasoning in the class form. The percentage figures in this table come from Ennis and Paulus (1965) and Roberge (1970). Of particular interest here are the variations in findings between researchers, despite the common conceptual framework, on certain principles. Principles 2,3, and 4 all show a wide fluctuation, but it is particularly noteworthy in Principle 3, denying the antecedent. The explanation for such wide fluctuations is not entirely clear, but it must be noted that the results reported for conditional reasoning are much more consistent across researchers. This is a matter that deserves further investigation.

The significant main effects differences in principles and types of reasoning confirm other researchers' (Ennis & Paulus, 1965; Gardiner, 1965; Paulus, 1967; Roberge, 1970; Roberge & Paulus, 1971) findings. The mastery data found in Table 7 perhaps has the greatest implications for teacher training, particularly if we take seriously other researchers' suggestions that their findings indicate that implicit instruction in logical skills can begin as early as grade 4. Clearly, large numbers of beginning elementary school teachers are deficient in these rudimentary logical skills of a judgment type.

TABLE 9
PERCENTAGES OF STUDENTS WHO HAD MASTERED EACH OF THE FIVE PRINCIPLES OF DEDUCTIVE REASONING IN CLASS FORM

Grade Level	Principle					n/N	p	Researcher
	1	2	3	4	5			
4	56	5	5	11	34	5/6	.02	Ennis & Paulus (1965)
4	60	2	0	44	35	8/12	.02	Roberge (1970)
6	79	15	15	17	44	5/6	.02	Ennis & Paulus
6	58	4	2	44	28	8/12	.02	Roberge
8	91	21	13	24	55	5/6	.02	Ennis & Paulus
8	82	5	4	79	74	8/12	.02	Roberge
10	92	44	35	37	80	5/6	.02	Ennis & Paulus
10	95	40	21	75	79	8/12	.02	Roberge
12	97	68	47	58	75	5/6	.02	Ennis & Paulus
Pre-service Elem. Tchrs. Jrs. & Srs.	91	19	33	27	65	6/8	.02	Jansson

Efficient instruction in these skills will require further investigation into the instruction of the various principles. Perhaps instruction in one or two of the principles, with appropriate techniques, will be sufficient for mastery of all five. The present study has dealt only with a judgment form, but an extension of the Paulus (1967) study could be made to include assumption finding, first in a situation of logical necessity and later in a more general situation.

The answers to the first three questions of this study have been made in the affirmative, while the results relative to the fourth question reveal the necessity of specific instruction for teachers—maturation alone does not provide the desired skills.

References

- Ennis, R. H. The development of a critical thinking test. Unpublished doctoral dissertation, University of Illinois, 1959.
- Ennis, R. H. A concept of critical thinking. *Harvard Educational Review*, 1962, 32 (Winter), 81-111.
- Ennis, R. H., & Paulus, D. H. *Critical thinking readiness in grades 1-12 (Phase I: Deductive reasoning in adolescence)*. Ithaca, N.Y.: The Cornell Critical Thinking Project, 1965.
- Gardiner, W. L. An investigation of understanding of the meaning of the logical operators in propositional reasoning. Unpublished doctoral dissertation, Cornell University, 1965.
- Hill, S. A. A study of logical abilities in children. Unpublished doctoral dissertation, Stanford University, 1961.
- Howell, E. N. Recognition of selected inference patterns by secondary school mathematics students. Unpublished doctoral dissertation, University of Wisconsin, 1966.
- Jansson, L. C. The development of an instrument to assess critical thinking ability in mathematics. Unpublished doctoral dissertation, Temple University, 1970.
- O'Brien, T. C., & Shapiro, B. J. The development of logical thinking in children. *American Educational Research Journal*, 1968, 5, 531-542.
- O'Brien, T. C., Shapiro, B. J., & Reali, N. C. Logical thinking—Language and context. *Educational Studies in Mathematics* 1971, 4, 201-219.
- Paulus, D. H. A study of children's abilities to deduce and to judge deductions. Unpublished doctoral dissertation, Cornell University, 1967.
- Roberge, J. J. A study of children's abilities to reason with basic principles of deductive reasoning. *American Educational Research Journal*, 1970, 7, 583-596.
- Roberge, J. J. Recent research on the development of children's comprehension of deductive reasoning schemes. *School Science and Mathematics*, 1972, 72, 197-200.
- Roberge, J. J., & Paulus, D. H. Developmental patterns for children's class and conditional reasoning abilities. *Developmental Psychology*, 1971, 4, 191-200.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

MARGARET M. CLIFFORD

University of Iowa

Validity of Expectation—A Developmental Function¹

Validity of self-expectations was found to be an increasing monotonic function of development among first-, third-, and fifth-graders. The relatively high proportion of invalid expectations observed at all grades is attributed to overestimation tendencies and an apparent failure to process task feedback. Given that expectation is a major component of human motivation theories, an investigation of the determinants of valid expectations and means by which they can be manipulated would seem desirable. (Dr. Clifford is Associate Professor in the Division of Educational Psychology, Measurement and Statistics, College of Education, The University of Iowa.)

It has been demonstrated that self-expectation is a predictor of task persistence (Battle, 1965; Feather, 1963) as well as academic achievement (Battle, 1965; Battle, 1966; Biggs & Tinsley, 1970; Keefer, 1969). There is also evidence that expectations affect the tendency to approach tasks (Crandall, Katkovsky & Preston, 1962; Feather, 1966) and solve problems (Feather, 1966; Tyler, 1958). Studies supporting these findings appear to be based on the assumption that individuals generally form relatively accurate expectations. However, there is reason to question the accuracy of children's expectations. They have been found to be affected by adult reactions (Crandall, 1963; Crandall, Good, & Crandall, 1964); it has also been demonstrated that children have a tendency toward wishful, and/or unrealistic expectations (Bialer, 1961; Irwin, 1944; Irwin & Mintzer, 1942; and Marks, 1951). A comparison of related studies by Marks (1951) and Irwin (1953) using similar procedures and measures further suggests that the tendency to overestimate one's ability decreases with an increase in age.

The quality of an expectation can be assessed from at least two perspectives—the conformity of an expectation to subsequent relevant performance and the conformity of an expectation to previous relevant

¹ This research was supported by Public Health Service Grant MH-24091-01. The cooperation of administrators, teachers and pupils from the Clear Creek and Washington Community School Districts in Iowa is gratefully acknowledged.

behaviors. The former concept, here referred to as the Predictive-Value of Expectation, is assumed to be dependent upon the latter concept, subsequently referred to as Validity of Expectation, which is of primary concern in this study.

It is predicted that, among children, Validity of Expectation is an increasing monotonic function of development and that invalid expectations are reflected to a greater extent in overestimation than in underestimation. It is also postulated that accuracy in processing feedback may partially account for Validity of Expectation.

Method

Subjects

The Clear Creek and Washington, Iowa school systems participated in this study. Observations were obtained from four classrooms at each of the first, third, and fifth grade levels. The number of males at each successive level was 47, 42, 49; the number of females was 47, 51, 41. Mean standardized achievement scores from each grade indicated the subjects were representative of average, middle-class students.

Materials

Two types of perceptual judgment problems were designed for this experiment—"string" problems (ST), in which the larger of two rectangular line drawings was to be identified, and "land" problems (LA), in which the larger of two irregularly-shaped areas was to be identified. A spot was located beside each of the members of the various pairs of drawings to be compared. Supposedly, the spot corresponding to a "correct response" was completely erasable, whereas that associated with an "incorrect response" was only partially erasable. In reality, the line lengths or areas were equal, and both spots for any given problem were either erasable or non-erasable.

Three pages, each containing 10 land problems, and another three pages, each containing 10 string problems, were prepared; pages will be identified with numerical subscripts. Four of these six sheets were then assembled in one of two sequences (LA₁, ST₁, LA₂, ST₂ or ST₁, LA₁, ST₂, LA₂). For half of the 4-sheet packets in each sequence the erasable and non-erasable spots were arranged to convey "correct" responses totalling 8, 3, 7, 2, on the four successive pages. The manipulated feedback on the other half of the booklets was 3, 8, 2, 7. In any given booklet, therefore, one type of problem was associated with high feedback while the other was associated with low feedback (*e.g.*, the scores on the land pages were 3 and 2 while the scores on the string pages were 8 and 7). Type of task and level of feedback were counterbalanced. A fifth page, which was either LA₃ or ST₃, completed the booklet. It was used only to elicit a subject's expectation, and was not worked; therefore, no feedback was obtained on this final sheet.

On the back of the booklet, three follow-up items were provided. The subject was asked to indicate (a) whether he found the string or land problems easier, (b) the number of correct responses he made on page one, and (c) whether an example problem, supposedly worked by another student, was "right" or "wrong"; this example depicted a land problem in which an attempted erasure resulted in failure to remove the entire spot (*i.e.*, an incorrect response).

Procedure

Data collection was conducted in intact classrooms to insure a familiar achievement-oriented setting and subjects' typical attentiveness to academic feedback. The compilation of booklets involved a counterbalancing procedure which yielded 16 different booklet forms. A systematic distribution of these forms insured that no two adjacent subjects worked similar pages at the same time. The task was explained with the use of two solvable sample problems, keyed to reflect objectively correct and incorrect responses. Subjects were told the problems in the booklet would be more difficult than the examples, but that they should do their best. Allowing about 1½-2 minutes per page, the experimenter directed the subjects to work the first four pages. An emphasis on a limited time allowance and use of a stopwatch served to economize on time and to minimize student idleness as well as the opportunity for peer comparisons. However, subjects were given time to complete each page.

Upon turning to page 5, subjects were asked to indicate in writing at the top of the page how many of the 10 problems on that page they thought they would answer correctly. Without working the last page, subjects were immediately assisted in reading the follow-up items. An informal debriefing, consisting of answering subjects' questions, was conducted; the experimenter explained that it was not necessary to work the last page, since everyone had worked at least two pages of each type and time was running out.

Expectation Measures

The measures of expectation were the Raw Score Expectation (E_r) recorded on page 5 of the booklet and the Validity of Expectation (E_v) which was a measure of the discrepancy between E_r and the manipulated scores on the two preceding pages of corresponding type (*e.g.*, if page 5 consisted of LA_3 , the mean feedback on LA_1 and LA_2 was compared with the expectation recorded by subject on page 5). The scores obtained on the pages similar in type to page 5 are subsequently referred to as "relevant feedback."

Validity of Expectation was coded as follows: A score of 1 was assigned to a subject whose Raw Score Expectation was two or more points lower than his lower relevant feedback (*i.e.*, E_r of 5 or less for subjects with relevant feedback of 7 and 8, and 0 for subjects with relevant feedback of 2 and 3) or 2 or more points higher than the higher relevant feedback (*i.e.*, E_r of 10 for subjects with relevant feedback of 7 and 8, and E_r of 5 or more for subjects with relevant feedback of 2 and 3). A score of 2 was assigned to a subject whose expectation was one point lower than his lower relevant feedback or one point higher than his higher relevant feedback. A score of 3 was assigned to a subject whose expectation coincided with either of his two relevant feedback scores (*i.e.*, 7 or 8 in the high and 2 or 3 in the low feedback condition).

Results²

A Sex x Grade analysis of variance of Validity of Expectation yielded a Grade effect, $F(2, 271) = 4.74$; $p < .01$. Means indicated an increase in Validity of Expectation with grade; they were 1.34, 1.41, 1.67 for first, third, and fifth respectively, and thus offered support for the prediction that there is an increasing monotonic relationship between these variables. However, Scheffé tests between adjacent grades were not significant at the .05 level. Neither the Sex effect nor the Grade x Sex interaction was significant.

A Sex x Grade analysis of variance of Raw Score Expectation also resulted in a Grade effect, $F(2, 271) = 10.77$; $p < .001$, in the absence of a Sex or interaction effect. The nature of this effect is, as expected, the reverse of that found for Validity of Expectation; means were 7.78, 7.04, and 6.30 for first, third and fifth respectively. Whereas Validity of Expectation increases with Grade, Raw Score Expectation decreases. Given that the average relevant feedback across subjects in any grade was 5, and the mean Raw Score Expectation exceeds 6 at every grade level, these data offer support for the second hypothesis that invalid expectations are reflected more in overestimations than in underestimations.

Once again, Scheffé tests between adjacent grades failed to indicate significant differences. However, the frequency with which subjects stated an expected score of 10 on page 5 supported the predicted developmental trend. This maximum score (unobtained by any subject on previous trials) was recorded by 44% of the subjects in the first grade, 20% in the third grade, and 3% in the fifth grade. The procedure for testing proportions reported in Blommers and Lindquist (1960) indicated all pairwise comparisons between these percentages were significant beyond the .001 level.

Additional analyses of variance were performed to investigate whether these data were dependent upon subjects stating an expectancy on a task associated with "success" vs "failure"; half the subjects predicted for a task on which they had received scores of 2 and 3 (Low Prediction Page), whereas half predicted for a task on which they had previously scored 7 and 8 (High Prediction Page). Two 2 x 3 (Prediction Page by Grade) analyses of variance showed that aside from the Grade effect previously reported on both Validity of Expectation and Raw Score Expectation measures, only one significant effect resulted; namely, Validity of Expectation was higher in the High as opposed to the Low Prediction Page condition. Given the pervading tendency to overestimate, this latter finding may reflect nothing more than a ceiling effect; the discrepancy between the maximum possible score and High relevant feedback cannot be as great as the discrepancy between the maximum possible score and Low relevant feedback.

A more detailed assessment of subjects' expectations under these two conditions is presented in Table 1. It suggests that High expectations of fifth-graders, unlike those of first- and third-graders, were associated almost exclusively with failure-oriented tasks. Using the procedure reported in Blommers and Lindquist (1960), the proportion of fifth-graders who recorded expectations of 9 or 10 was significantly greater ($p < .002$) for the

² The univariate analyses reported represent a least squares solution which allows for disproportionate cell frequencies.

TABLE 1
PERCENTAGE OF LOW, ACCURATE, AND HIGH EXPECTATIONS*
GIVEN BY THE HIGH AND LOW PREDICTION PAGE SUBJECTS

Prediction Page Type	Grade	N	Expectation		
			Low	Accurate	High
High	1	48	31	25	44
Low	1	46	04	00	96
High	3	45	44	22	33
Low	3	48	00	04	96
High	5	46	50	43	07
Low	5	44	00	05	95

* Low: below low relevant feedback
Accurate: equal to relevant feedback
High: above high relevant feedback

| pairwise test, $p < .05$ || pairwise test, $p < .01$

Low as opposed to the High Prediction Page. Comparable tests of proportions for first- and third-graders showed nonsignificant differences.

Table 1 also indicates that accurate expectations occurred almost exclusively in the High as opposed to the Low Prediction Page. Furthermore, whereas inaccurate responses in the High Prediction Page condition were rather evenly distributed between Low and High expectations for first- and third-graders there was a noticeable tendency for fifth-graders to express low more frequently than high expectations in this condition, 50% *vs* 7%, $p < .01$. (Several obviously significant contrasts, assumed to be a phenomena of overestimation or ceiling effects, which were discussed earlier have not been indicated in this table.)

A summary of the accuracy with which subjects responded to the three follow-up items is presented in Table 2. Whereas the first and third items allowed for only one of two responses (correct or incorrect), the second item allowed for one of 10 responses. A response to this item was considered accurate if the number of correct responses reported by the subject was no

TABLE 2
PERCENT OF SUBJECTS WITH ACCURATE RESPONSES ON FOLLOW-UP ITEMS

Items	Grade		
	1	3	5
Task Difficulty	52	67	80
Correct Responses	65	86	89
Judgment of Example	82	81	89

| pairwise test, $p < .05$ || pairwise test, $p < .01$

more than one point discrepant from the manipulated feedback on his first page. (This one-point discrepancy margin was in acknowledgment of the possibility of inadvertent errors in counting.) Any other response to this item was classified as inaccurate.

The increase in accuracy with increase in grade level found for item 1 lends credence to the speculation that accuracy in processing feedback may be one explanation for the developmental function of Validity of Expectation. The accuracy pattern for item 2 (assumed to require less complex skills than item 1) suggests that accuracy develops more rapidly with age, the less complex the feedback-processing task. The relatively high accuracy on item 3 and the consistency across developmental levels suggests most subjects could, indeed, discriminate between correct and incorrect feedback. Whereas third- and fifth-graders showed no difference in accuracy regardless of whether they interpreted their responses (item 2) or another's response (item 3), first-graders showed significantly greater accuracy in interpreting the feedback on the example as compared to their own first page. Furthermore, 85% of the first-graders who responded inaccurately on item two reported an inflated score for themselves.

Discussion

The developmental function demonstrated in this experiment suggests that if, as frequently assumed, valid expectations are a major component of human motivation, the manipulation of Validity of Expectation in children may warrant attention. While age emerges as the most obvious factor accounting for variability of Validity of Expectation, it obviously precludes manipulation. If, however, one postulates that this developmental function of Validity of Expectation can alternatively be explained by an increasing monotonic relationship between Age and accuracy in processing feedback, training on skills such as perception, recognition, discrimination, association, and integration would appear to be viable techniques by which to affect Validity of Expectation. More specifically, it is postulated that providing concrete, visual feedback (e.g., charts, tallies) on performance, will increase Validity of Expectation. Similarly self-scoring and recording of performance may also enhance Validity of Expectation for children.

Furthermore, if one postulates that this developmental function of Validity of Expectation is partially explained by a decreasing monotonic function between age and an overestimation tendency, reality training, (e.g., emphasizing the difference between what one would like to have happen and what one thinks will happen in a given situation) may provide a means of manipulating Validity of Expectation. As suggested in the results, the responses on the follow-up items seem to lend credence to these postulations. Finally, given that even for fifth graders Validity of Expectation, as coded in this study, does not exceed 1.8 on a 3-point scale, it seems advisable to reinvestigate this variable on an extended developmental continuum.

Validity of Expectation—A Developmental Function

References

- Battle, E. Motivational determinants of academic task persistence. *Journal of Personality and Social Psychology*, 1965, 2, 209-218.
- Battle, E. Motivational determinants of academic competence. *Journal of Personality and Social Psychology*, 1966, 4, 634-642.
- Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. *Journal of Personality*, 1961, 29, 303-320.
- Biggs, D. A., & Tinsley, D. J. Student-made academic predictions. *Journal of Educational Research*, 1970, 63, 195-197.
- Blommers, P., & Lindquist, E. F. *Elementary statistical methods*. Houghton-Mifflin Company, Boston, 1960.
- Crandall, V. C. Reinforcement effects of adult reactions and nonreactions on children's achievement expectations. *Child Development*, 1963, 34, 335-354.
- Crandall, V. C., Good, S., & Crandall, V. J. Reinforcement effects of adult reactions and nonreactions on children's achievement expectations: A replication study. *Child Development*, 1964, 35, 485-497.
- Crandall, V. J., Katkovsky, W., & Preston, A. Motivational and ability determinants of young children's intellectual achievement behaviors. *Child Development*, 1962, 33, 643-661.
- Feather, N. T. Persistence at a difficult task with alternate task of intermediate difficulty. *Journal of Abnormal and Social Psychology*, 1963, 66, 604-609.
- Feather, N. T. Effects of prior success and failure on expectations of success and subsequent performance. *Journal of Personality and Social Psychology*, 1966, 3, 287-298.
- Irwin, F. W., & Mintzer, M. G. Effect of differences in instructions and motivation upon measures of the level of aspiration. *American Journal of Psychology*, 1942, 55, 400-416.
- Irwin, F. W. The realism of expectations. *Psychological Review*, 1944, 51, 120-126.
- Irwin, F. W. Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality*, 1953, 21, 329-335.
- Keefer, K. E. Self-prediction of academic achievement by college students. *Journal of Educational Research*, 1969, 63, 53-56.
- Marks, R. W. The effect of probability, desirability, and privilege on the stated expectations of children. *Journal of Personality*, 1951, 19, 332-351.
- Tyler, B. B. Expectancy for eventual success as a factor in problem solving behavior. *Journal of Educational Psychology*, 1958, 49, 166-172.

LOIS FOSTER
and
MARY NIXON
University of Alberta

The Interview Reassessed

In this paper the interview is discussed as problematic. This approach leads to an examination of the interview as "a product of social arrangements" (Young, 1972, p. 194) rather than a prescription for strategies of operation. Such discussion may raise new questions with the possibility of new answers. (Lois Foster and Mary Nixon are graduate students in the Faculty of Education, The University of Alberta.)

The recent trend in sociology of education which has attacked the positivistic stance of much of the reported research has been accompanied by a renewal of interest in other approaches (Young, 1971; Hopper, 1971; Filmer *et al.*, 1972). As a result there has been an increasing emphasis on research based on the interpretative paradigm which challenges the dominance of highly statistical research oriented to the normative paradigm (Colfax & Roach, 1971; Brown, 1973). The dominant theory underlying the normative paradigm is structural functionalism which seeks an understanding of a whole system by analysing the function served by each of the individual parts. The interpretative paradigm derives its orientation from such theoretical bases as symbolic interactionism and phenomenology which focus upon the interactions of people. For the researcher, working within the normative perspective, objectivity as inter-subjective testability is a major goal. In contrast, the interpretative paradigm takes account of the interaction between researcher and researched and incorporates it within the research design.

At the present time there is a tendency to look upon these two paradigms as discrete. Yet a careful appraisal of the positivistic work done by such sociologists as Halsey, Floud and Anderson (1961) reveals an appreciation for the underlying questions raised by the interpretative approach (Eggleson, 1973). This would suggest the possibility, at least, that these two approaches need not be mutually exclusive.

The interview is probably the most adaptive instrument of data collec-

tion. The researcher may place the emphasis deliberately upon the interaction between researcher and participants or the emphasis may be placed on the elimination of reactive effects by the use of *a priori* hypotheses and closed response categories. An alternative to this dichotomy is a format which combines predetermined questions designed to elicit specific information with spontaneous questions that arise from the interaction between the researcher and the participant. Surprisingly, the interview method, although it is as much the result of research orientation as the most structured, written survey technique, does not appear to have come in for as much criticism as the questionnaire method.

Opportunities for some experience in diverse interview situations have suggested the need to examine the interview as problematic. To define the interview as problematic is to define it as "a product of social arrangements," (Young, 1972, p. 194) that is, as a result of the activities of people. Textbooks such as Selltiz *et al.* (1965), Phillips (1966), Good (1966), treat the interview as a conventional typification comprised of "taken for granted" definitions of phenomena which help to make the everyday world meaningful to the actor (Schutz, 1962). The assumption of a common definition or a common research purpose may not be warranted. Methodology courses are frequently little better with their suggestion that the interview is the "human" replacement of the written questionnaire. The implied criterion for the use of the interview is often the extent of the funds available to the researcher. Those studies which have substantial funding may choose to use the interview, but as an addition to an extensive printed survey. Modestly funded studies may select the interview, particularly the telephone interview, as an inexpensive method of data collection. Another criterion for the use of the interview is to improve an unacceptably low response rate obtained by a written questionnaire.

Textbook writers tend toward an atheoretical treatment of the interview. Indeed it is often dismissed. For example, one recent text on survey and research methods devotes fourteen out of a total of three hundred and seventy-two pages to general rules for interviewing (Babbie, 1973, pp. 171-185). Such information as "the interviewer should be relaxed and friendly without being too clinging" (Babbie, 1973, p. 173), that he should be familiar with the questions (Babbie, 1973, p. 174), or that the time of day should be carefully controlled (Travers, 1969, p. 236), are practical hints of a mundane nature. Less helpful is the explanation that "the term 'interview' is constructed out of two words, 'inter' and 'view'" (Lofland, 1971, p. 75), while lack of solutions detracts from the utility of Glazer's warnings of the problems to be encountered in interview research (Glazer, 1972, p. 25). Those who seek a theoretical treatment of the interview are likely to be disappointed.

It is not always recognized that the structured interview differs only in format from the structured questionnaire. In this type of research, meaning is imposed on the situation by the questions asked and by the researcher's attitude which filters out information that is seen as peripheral to the phenomenon under investigation. The problematic approach would raise questions of (a) the influence of value and belief systems, (b) ways of defining legitimate data, and (c) the emphasis and direction of the research report.

The adoption of a problematic approach stimulates a rethinking of the interview, the justification for its use and the legitimacy of the data outcomes. Interview research ranging from children without choice as to their participation to influentials who volunteered their resources of time and knowledge provided the experiential base. Yet this diversity lent itself to a classification on two dimensions: influential . . . non-influential and selected . . . self-selected. The definitions of influential . . . non-influential and selected . . . self-selected are not mutually exclusive categories since an individual who can be termed "self-selected" in one research situation may be constrained to participate in another, thus becoming, in this sense "selected." In general, individuals or groups may choose whether they respond to the researcher's request that they participate in a study but for certain groups (*e.g.* children) and in certain situations (*e.g.* government sponsored studies) the decision to participate is not voluntary. Societal status has an important bearing not only on an individual's perception of his position relative to the implications of the study but also on the researcher's perceptions which affect the design of the study.

Recognition of the importance of values, those of the researcher and the researched, both influentials and non-influentials, transforms assumptions into questions. The assumption that reactive effects are necessarily sources of error appears in such authoritative texts as that of Rundle and McGrath (1972, p. 184). A design attempting to reduce these "sources of error" will tend toward a format which emphasizes prescribed wording. A positive view of reactive effects would contest this assumption. It raises questions as to qualitative ways of knowing. The researcher faces a decision as to whether he will attempt a more complex design that will incorporate qualitative data or whether he will restrict himself to quantifiable data which can be manipulated with appropriate statistical tools.

Even if the researcher is content with quantifiable data he cannot with impunity ignore the participants' value systems since values and beliefs influence the responses. It would indeed be a dangerous assumption that there can exist a question that, for all individuals and groups, could be entirely value free. A structured format, however, minimizes the intrusion of values and enhances a study's replicability. The interpretative approach to research design takes into account qualitative data but its objectivity may be more readily challenged.

The researcher, influenced by the positivistic approach, may too readily make the assumptions: (1) that a sophisticated statistical treatment of his data leads to a sophisticated understanding of his topic, and (2) that questions will elicit answers reflecting a participant's "true" position on the concern of the study, that is, that responses that are reliable over time are indeed valid. Those not committed to a normative approach would suggest: (1) that a qualitative treatment of data is a valuable contribution to understanding, and (2) that the participants' perceptions become the concern of the study. In practice the first approach will tend toward an interview situation where the participant is required to give answers structured to allow categorization of responses suitable for transfer to computer cards. The second approach is more open-ended and thus the data that is gathered frequently eludes a ready classification. Constraints are imposed on the methodology by the researcher's perception of the influential or non-

influential status of the interviewees. Legitimate data takes a variety of forms. In certain cases, for example, the collection of biographical data, the use of the second approach would be time consuming and probably unnecessary.

The reports which result from these two perspectives will differ in emphasis and direction but both are subject to bias. The emphasis of normative research is on *a priori* hypotheses which are to be supported or rejected by statistical data analysis. This imposes defined limits on the research report because once the findings have been presented the conclusions are almost self-evident. However, shifting cutting points and levels of significance manipulate the findings in the direction of the researcher's bias. The interpretative approach makes no formal statement of hypotheses in advance and therefore the analysis of the data presented in the report is subject to greater variation. Recognizing that bias cannot be eliminated, the researcher attempts to analyze his own position in relation to the data. Thus intuitive rather than quantifiable ways of knowing take precedence in this approach.

Reassessment of the interview arose from the variety of definitions derived from diverse interview situations. Considering the interview as problematic permits a scrutiny of assumptions and strategies which are generally left unquestioned. Yet, as of the present, the interpretative approach lacks operational definitions of concepts. Interviews conducted from this perspective would be different from, but not necessarily better than, those conducted within the normative framework. With more examples of the use of the interpretative approach, the result may be that the two perspectives are profitably incorporated into a new research paradigm.

References

- Babbie, E. *Survey research methods*. California: Wadsworth Publishing Co., Inc., 1973.
- Brown, R. (Ed.) *Knowledge, education and culture*. London: Tavistock Publishing Co., 1973.
- Colfax, T., & Roach, J. (Eds.) *Radical sociology*. New York: Basic Books, 1971.
- Eggleston, S. *Sociology and the sociology of education*. University of Alberta, November, 1973. (Mimeo.).
- Filmer, P., Phillipson, M., Silverman, D., & Walsh, B. *New directions in sociological theory*. London: Collier-Macmillan Publishing Co., 1972.
- Halsey, A., Floud, J., & Anderson, C. *Education, economy and society*. Glencoe, Ill.: The Free Press, 1961.
- Glazer, M. *The research adventure*. New York: Random House, 1972.
- Good, C. V. *Essentials of educational research, methodology and design*. New York: Appleton-Century Crofts, 1966.
- Hopper, E. (Ed.) *Readings in the theory of educational systems*. London: Hutchinson, 1971.
- Lofland, J. *Analyzing social settings*. California: Belmont Publishing Company, 1971.
- Phillips, B. *Social research: Strategy and tactics*. New York: The Macmillan Co., 1966.
- Rundle, P., & McGrath, J. *Research on human behaviour*. New York: Holt, Rinehart and Winston Publishing Company, 1972.

- Schutz, A. *Collected papers Volume 1: The problem of social reality*. M. Natanson (Ed.) The Hague: Martinus Nijhoff, 1962.
- Selltiz, C., Jahoda, M., Deutsch, M., & Cook, S. *Research methods in social relations*. New York: Holt, Rinehart and Winston, 1964.
- Travers, R. *An introduction to educational research*. New York: The Macmillan Company, 1969.
- Young, M. F. D. *Knowledge and control*. London: Collier-Macmillan Publishing Company, 1971.

V. FROESE

The University of Manitoba

Does The Dolch Do?

Considerable controversy has been generated recently with regards to the usefulness of various word lists. The present study deals with students' actual recognition of words on the Dolch 220 list and another comparable list drawn from the Kučera/Francis corpus.

The Dolch 220 appears to differentiate between grades and IQ levels and correlated more highly with a criterion measure administered five months later. Both tests did not differentiate between the sexes or on socioeconomic status. (V. Froese is Assistant Professor in the Department of Curriculum, Faculty of Education, The University of Manitoba.)

Considerable interest has been generated recently in word lists by Johnson (1971, pp. 449, 451), Johns (1971), Otto and Chester (1972), Durr (1973), Harris and Jacobson (1973-74) and Pennock (1974). Consequently, the inveterate Dolch 220 has come under close scrutiny. Although the pervasiveness of this list is not precisely known, Jacobs (1967, p. 2) concluded from his interview of teachers that (1) the Dolch lists are familiar to most teachers in remedial and primary reading, (2) the use of the lists may be greatest in the area of remedial reading, and (3) few people are aware of the origins and ages of the lists.

The Dolch 220 also known as *The Basic Sight Word Test*, was based on vocabulary studies of the 1920's and is still used in its original form. The current form was copyrighted by Garrard Publishing Company in 1942.

It seems logical to question the usefulness of this list simply because of its age and it is not surprising to find comments such as the following:

The Dolch List, as a corpus, has outlived its usefulness and . . . a more adequate substitute is available (Johnson, 1971, pp. 449, 451).

Yet the fact is that in addition to becoming somewhat passé over more than three decades, the list [Dolch 220] had a questionable empirical base in the first place (Otto & Chester, 1972, p. 435).

This paper describes a study which was designed to investigate the current usefulness of the Dolch 220. It was decided to test second and third grade children's actual recognition of the Dolch words and compare the

results to scores on a test suggested by Johnson (1971). It is important to note, then, that this study deals with *recognition* of words and is not a study based on text demand (frequency counts, occurrence in basal readers, etc.). To provide some indication of the test's predictive validity, the *Botel Word Opposites Test* (1966) was administered five months later. This test was used since it correlated highly, from 0.82 to 0.90, with other comprehension measures (Botel, 1970).

Before describing the results it is instructive to consider the research on word recognition tests. It is essential to keep in mind four factors influencing vocabulary study: (1) the source—whether from children or adults, (2) the modality—whether oral or written, (3) the method of collection—whether frequency counts, free-association, from basal readers, tested knowledge, adult judgment, or some combination, and (4) the period or time of collection—whether in the 1920's or late 1960's and early 1970's, etc.

Related Research

The importance of a basic sight vocabulary was clearly articulated by Dolch. His list contained the words which indicated the structural relationships between members of the form classes (nouns, verbs, adjectives, adverbs) as well as a number of "irregular" words, both being essential to all subject matter material. Furthermore, Dolch (1960) demonstrated that the basic sight vocabulary comprised over one-half of the running words in the subjects of Reading, Arithmetic, Geography, and History.

Johns (1971) replicated this analysis for five current reading series and found that the Dolch words still comprised over one-half of the running words, although the percentage at each grade level (1-6) was somewhat lower. Johnson and Barrett (1972) reported similar findings for Johnson's list which contains many of the same words.

The similarity of these counts appears to indicate that publishers take into account the various studies of basic vocabularies or that they are governed by the fact that these words are necessary to all writing.

Vocabulary research has also influenced the measurement of reading difficulty, since according to Klare (1963), vocabulary accounts for the greatest amount of variance in readability formulas. And since readability formulas are also utilized to control basal reader material, the cycle of interdependent factors pointed out by Froese (1971) is completed.

A further problem related to the study of vocabulary is an environmental one—changes in technology, advancements in the sciences, and cultural influences have their concomitant language influences. Jacobs (1967) has documented one aspect of this change by replicating the Buckingham-Dolch Free-Association Word Study. He concluded that the 1926 Free-Association List is no longer representative of current students' vocabularies in terms of list content, grade-level assignment, and student performance.

Johnson (1971, p. 451) also referred to this change when he stated: "Of the 220 words on the Dolch Primary Word List, representative of the 1920's, *eighty-two* words or *37 percent* are *not* among the most frequently occurring 220 words in the Kučera and Francis corpus compiled in the 1960's." This statement must, however, be interpreted carefully since the Kučera and

Francis study is based entirely on frequency of occurrence, includes nouns, and is derived from adult material—all of these are facets not common to the Dolch 220.

Lexicography also is concerned with language shifts and regular revisions of dictionaries are practically assumed. The motivation for the *Word Frequency Book* (1971) rests on these grounds but breaks new ground in applying computerized techniques to the vocabulary of children in grades 3 through 9 (and was likely inspired by Kučera's work with adult vocabulary).

A British study by Edwards and Gibbon (1964) examined the written vocabulary of children ages six to eight years and found that it expressed "exciting scientific wonders," that their lists overlapped but were not closely alike the American lists (of Rinsland, Thorndike, Gates and Dolch), and that the earlier lists were "out-of-date in respect of the scientific advances of recent years."

A further consideration to be taken into account when studying vocabulary is the shortcoming of frequency counts alone. Familiarity, meaning, and comprehension do not bear a one-to-one relationship to frequency as pointed out by Klare (1963). Dolch (1951) noted this and suggested that the factors of opportunity of use and emotional set are the causes of the discrepancy between frequency of use and tested word knowledge.

The Study

The purpose of this study was to compare students' tested recognition of the Dolch 220 words with their responses to the 200 highest frequency words found by Kučera and Francis in their *Computational Analysis of Present-Day American English* (1967). The scores on the Dolch test were also compared to current student I.Q.'s and comprehension scores as tested five months later by the *Botel Word Opposites Test*.

Subjects for the study were 155 second graders and 179 third graders from classrooms in four schools in a large suburban school division in Winnipeg, Manitoba, Canada. Winnipeg is a city with a population of approximately 560,000. The factors of grade level, I.Q. level, socioeconomic level, and sex were considered.

The findings are presented in Table 1.

From a statistical analysis, it was found that significant differences occurred at the .05 level between tests, between grades, and among I.Q. levels on the Dolch test. On the Kučera/Francis 220 only the difference between middle and lower I.Q.'s was significant. The differences between socioeconomic levels and sexes did not reach significance.

A further analysis consisted of correlating the I.Q.'s, *Botel Word Opposites* scores, and the scores on the two vocabulary test. The results are presented in Table 2.

All correlation coefficients except those marked NS (non-significant) are significant at the .01 level.

The fact that few grade three students made errors on these tests obviously resulted in a ceiling effect and hence the non-significant findings.

However, the correlations between the Dolch 220 and I.Q., as well as

TABLE 1
MEANS AND STANDARD DEVIATIONS OF SELECTED VARIABLES
ON THE DOLCH 220 and the KUČERA/FRANCIS 220
HIGHEST FREQUENCY WORDS

Variable	Subjects	Means and Standard Deviations	
		Dolch 220	Kučera/Francis 220
Grade			
2	155	205 (19.5)	211 (17.7)
3	179	217 (5.1)	218 (3.8)
I.Q.			
Upper	118	216 (7.6)	218 (5.5)
Middle	151	212 (14.6)	216 (11.9)
Lower	66	203 (21.2)	209 (20.4)
SES			
Middle	181	213 (12.2)	216 (8.3)
Lower	154	210 (17.7)	214 (16.7)
Sex			
Boys	164	210 (16.0)	215 (11.6)
Girls	170	213 (14.0)	216 (14.2)

TABLE 2
INTERCORRELATION FOR THE DOLCH 220, THE KUČERA/FRANCIS 220,
THE BOTEL WORD OPPOSITES TEST AND OTIS I.Q.'S

	Dolch 220	Kučera/Francis 220
Grade 2		
I.Q.	.38	.30
Botel	.51	.42
Kučera/Francis	.84	-
Grade 3		
I.Q.	.09 NS	.01 NS
Botel	.38	.20 NS
Kučera/Francis	.82	-

between the Dolch 220 and the delayed Botel test are considerably higher than the same comparison for the Kučera/Francis 220.

Conclusions

A consideration of the data leads to the following conclusions about the Dolch 220:

1. The Dolch 220 test appears to differentiate between the grade levels considered.
2. The Dolch 220 test appears to differentiate between all three I.Q. levels considered.
3. The Dolch 220 test correlates higher with I.Q. scores than the other test considered.

4. The Dolch 220 test correlates higher with a comprehension measure administered five months later than does the other test considered.

In consideration of the original question, the conclusion appears to be that the Dolch does do!

This study was supported by the University of Manitoba Research Board.

References

- The Basic Sight Word Test*, Dolch, E. W. Champaign, Ill.: Garrard Publishing, 1942.
- Botel, M. *The Botel reading inventory*. Chicago: Follett Educational Corporation, 1966.
- Botel, M., Bradley, J., & Kashuba, M. The validity of informal reading testing. In W. K. Durr (Ed.), *Reading difficulties: Diagnosis, correction, and remediation*. Newark, Delaware: IRA, 1970, 85-103.
- Carroll, J. B., Davies, P., & Richman, B. (Eds.) *Word frequency book*. New York: Houghton-Mifflin and American Heritage, 1971.
- Dolch, E. W. Tested word knowledge vs. frequency counts. *Journal of Educational Research*, 1951, 44 (February), 457-470.
- Dolch, E. W. *Teaching primary reading*. (3rd ed.) Champaign, Ill.: Garrard Press, 1960, 258.
- The Dolch list re-examined—a further comment. In Crossfire, *The Reading Teacher*, 1972, 25 (March), 533.
- Durr, W. K. Computer study of high frequency words in popular trade juveniles. *The Reading Teacher*, 1973, 27 (October), 37-42.
- Edwards, R. P. A., & Gibbon, V. *Words children use: A survey of the words used by children in infants' schools with the resultant graded vocabulary*. London: Burke Publishing, 1964, 105.
- Froese, V. Word recognition tests: Are they useful beyond grade three? *The Reading Teacher*, 1971, 24 (February), 434.
- Harris, J. A., & Jacobson, M. D. Some comparisons between basic elementary reading vocabularies and other word lists. *Reading Research Quarterly*, 1973-1974, 9, 87-109.
- Jacobs, H. D. A free-association word list for the Willamette Valley. Bulletin of the Oregon School Study Council, Vol. II, University of Oregon, 1967.
- Johns, J. L. The Dolch Basic Word List—Then and now. *Journal of Reading Behavior*, 1971, 3 (Fall), 35-40.
- Johnson, D. D. The Dolch list re-examined. *The Reading Teacher*, 1971, 24 (February), 449, 451.
- Johnson, D. D., & Barrett, T. C. Johnson's basic vocabulary for beginning reading and current basal readers: Are they compatible? *Journal of Reading Behavior*, 1972, 4 (Fall), 1-11.
- Klare, G. D. *The measurement of readability*. Ames, Iowa: Iowa State University Press, 1963, 164-169.
- Kučera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, Rhode Island: Brown University Press, 1967.
- Otto, W., & Chester, R. Sight words for beginning readers. *The Journal of Educational Research*, 1972, 65 (July-August), 435-443.
- Pennock, C. D. Quick word lists for Canadian readers. *The Alberta Journal of Educational Research*, 1974, 20 (1), 8-13.

DOYAL NELSON

and

DAIYO SAWADA

University of Alberta

Studying Problem Solving Behavior in Young Children—Some Methodological Considerations

This is a report of certain aspects of a research project, "The Nature and Development of Problem Solving Behavior in Early Childhood." In the project, all raw data were recorded on video tape. Some considerations to be kept in mind when video tape is used in research are discussed. Some preliminary but tentative outcomes of the study are also reported. (Dr. Nelson is Professor, and Dr. Sawada is Associate Professor in the Department of Elementary Education, The University of Alberta.)

Mathematics instruction at the early childhood level should be aimed at helping children solve problems associated with situations and events they find themselves in from day-to-day. This problem solving cannot be the usual kind in which the solver seeks to find some mathematical expression to fit reality. The symbol system of the child is too incomplete to process problems of any consequence in this way. However, it should be possible to get children involved in solving problems which require the use of a variety of mathematical processes and which require the use of no symbols at all.

One difficulty with the assumption just stated is that not much is known about creating such problems for children and not much is known about the subsequent behavior of children in solving them. A research study is now under way at the University of Alberta, Department of Elementary Education, which addresses itself to these and other aspects of this matter.¹

The purpose of this paper is:

1. To discuss the guidelines used in the study to develop "good" problems for young children and to describe the nature of some problems so developed.

¹ This research was supported by a Canada Council research grant; Director: Doyal Nelson; Co-investigator: Daiyo Sawada.

2. To discuss the video tape recorder as a device for gathering data on the problem solving behavior of young children.

The Guidelines

Some years ago, a colleague, Dr. Joan Kirkpatrick, and the investigators set about to construct a model which could be used to create "good" problems for young children. We wanted to be reasonably sure that if the criteria stated in the model were adhered to, that the resulting problem or problem situation would indeed stimulate problem solving activity on the part of the child. The criteria which finally emerged as follows:

Criteria for Creating Good Problems

- A. *The problem should be of significance mathematically.* It is for the potential of the situation as a vehicle for the development of mathematical ideas that we choose a particular problem situation or family of situations above all others.
- B. *The situation in which the problem occurs should involve real objects or obvious simulations of real objects.* The main consideration here is that it be comprehensible to the child and easily related to his world of reality.
- C. *The problem situation should capture the interest of the child either because of the nature of the materials, the situation itself, the transformations the child can impose on the materials, or because of some combination of these factors.*
- D. *The problem should require the child to make moves or transformations or modifications with or in the materials.* It is difficult to overemphasize the role of action in early childhood learning. Most of the mathematical models we are interested in at this level are what might be called "action models."
- E. *Wherever possible, problems should be chosen which offer opportunities for different levels of solution.* If the child can move immediately from the problem situation to an expression of its mathematical structure it is not a problem.
- F. *Whatever situation is chosen as the particular vehicle for the problems, it should be possible to create other situations which have the same mathematical structure.* Another way of saying this is that the same problem should have many physical embodiments. It may not be possible for a child to generalize a solution to a certain kind of problem until the problem has come up in a variety of situations.
- G. *Finally, the child should be convinced he can solve the problem and should know when he has a solution for it.* If the child is somehow required to react with or transform materials used in problem situations it is usually easy to determine whether the problem meets the criterion or not.

The model is still a crude one but its application has been reasonably effective in providing direction for the creation of good problem situations for young children. Refinements are being made to the model as it is being used. The following section provides some samples of problems and problem situations which have resulted from the application of the

guidelines in the model. These were developed by a group of graduate students and staff over a short period of time.

Some Problem Situations

In this section, six problem situations are described as examples of the applications of the model criteria.

1. *Fold Out Shapes*. One important thing for children to learn is how three dimensional things can be represented in two dimensions. It is possible, of course, to disassemble and reassemble boxes of various shapes or to look at projections of these shapes or to trace around their faces. None of these activities, however, gives the child the control that our criteria suggest.

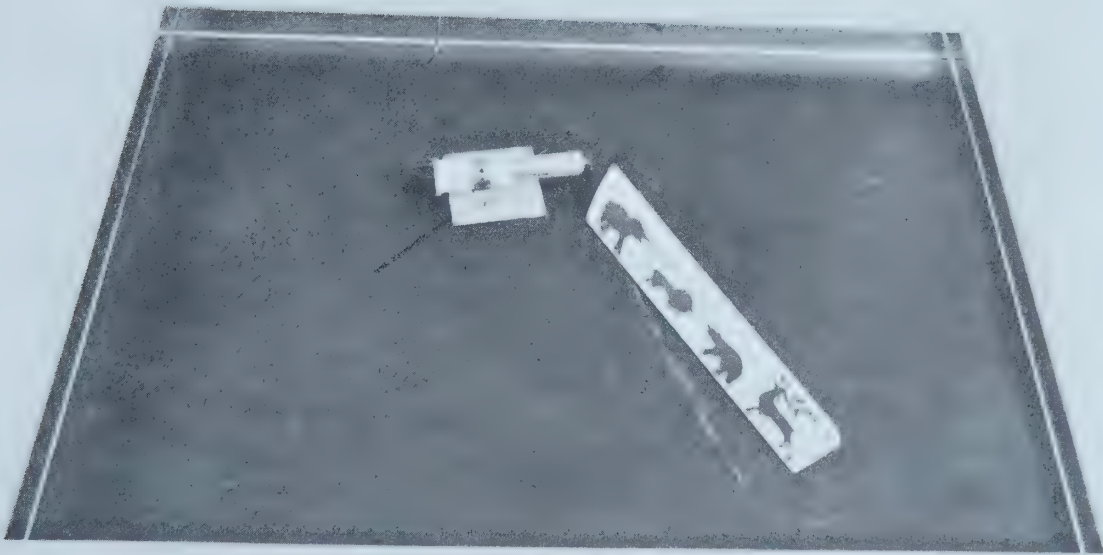
To overcome this difficulty, squares, rectangles, triangles, pentagons, and hexagons all with dimensions of ten centimeters or some multiple of ten centimeters were cut from large sheets of plexiglas. Then, on alternate edges, the positive (hooked) side of Velcro was glued. The negative (plush) side was glued to the remaining edges. (Of course, two positive or two negative faces would show up on one pair of adjacent edges in the triangles and pentagons.) Now the child could be provided with a set of faces which would make a solid join when put together but at the same time would be easy to take apart. A set of the Fold Out Shapes is shown in Photograph 1.



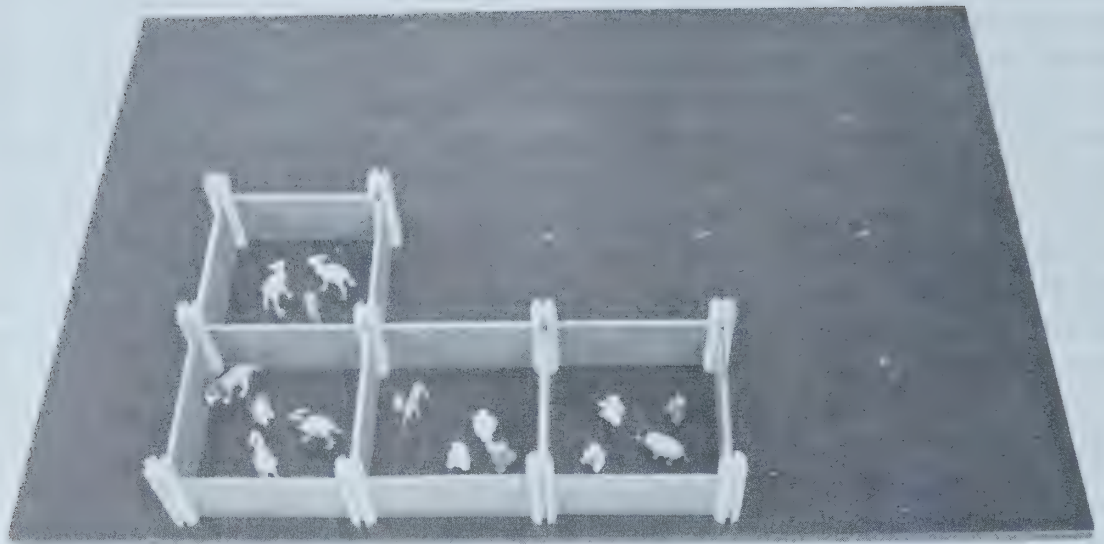
The kind of problems this material opens up for the child is very extensive. With sufficient pieces the Platonic solids can be constructed as well as a number of interesting, compound shapes.

2. *Mirror Reflection.* The next photograph shows the layout used to help children trace out the nature of a reflected beam of light. The board itself has shallow sides to which are attached mirrors. A point source of light was mounted on a carrier which could be moved along a groove and rotated 360° . There was a movable target which consisted of four animals. The basic problem was for the child to reflect the light off one mirror so that it fell on a designated animal. Then the child was asked to make moves with the source so that the beam fell on other animals.

Of course, the target could be moved in any position or the light could be reflected from two or even three of the mirrors. The child was asked to specify the path the light took to hit the target. The co-ordination of adjustments on the part of the child has been of primary interest in this problem.



3. *Animal Groups.* The ability of the child to make equal groups either by measurement or partition could be determined by this problem. The layout is shown in the photograph.



It consists of a large board in which holes are drilled to make a grid. Grooved dowels are cut down so that they fit in the holes of the board. The fit is quite tight so the dowels stand up rigidly. Panels are cut to fit the grooves in the dowels. These panels can be inserted by the child to make cages for the animals. The photograph indicates how the cages are assembled. One basic problem was to give the child a set of the animals and get him to make enough cages so that a given number of animals would go in each cage (measurement division). Another was to give the child a set of animals and a certain number of cages. He was asked to distribute the animals among the cages so that each cage would have the same number of animals (partitive division). Of course, a great variety of problems of this type can be devised around this particular apparatus.

4. *Theatre Grid.* In order to help children think about moving systematically in three dimensions, the theatre grid was developed as shown in the photograph below. The floors are each of a different color but the rows and seats in the row are indicated on each floor by a letter and a numeral. The child was then given a toy man and a ticket. The object was to find the correct seat for the man as indicated on the ticket. The color of the symbols on the ticket indicated which floor, the letter of the row, and the numeral of the seat in the row.



5. *Circular Sequences.* Problems associated with sequences can be designed around the carousel shown in the photograph.

The carousel has 12 basic compartments but each compartment can be provided with a partition to make 24 compartments in all. The front of the carousel can be adjusted from a closed position to positions which reveal one, two, or three compartments at a time. The compartments can be charged from the back in any kind of sequence with any kind of objects. The simplest would be to alternate two kinds of objects, for example, toy elephants and toy airplanes.

The child can be given information about the sequence by opening the front and turning the carousel. We are particularly interested in how much information a child needs to predict with certainty what the next object would be. A whole array of interesting problems come out of this apparatus.

6. *Factor Platform.* The factors of compound numbers can be illustrated physically by equal piles of objects. The factor platform shown in the accompanying photograph was devised so that the mechanics of piling would be minimized. A certain number of blocks were fitted in a number of grooves in piles of differing heights. The child then predicts whether the blocks can be rearranged so there will be the same number of blocks in each





row and verifies his prediction. Various strategies for testing the prediction can be observed among young children.

Basic Research Procedures

The research referred to in the early part of this paper is concerned with creating good problems, developing protocols for use with the problems, recording the behavior of children while engaged in the problems, and finally, analysing the problem solving behaviors.

To accomplish this purpose, six problems and a set of six different but mathematically equivalent problems were developed. A sample of ninety children was selected, fifteen from each age level ages three to eight. Each child was scheduled for a problem solving session. There was no time limit on the session, but the child was presented with six different problem situations in the session. Most children completed the session in less than an hour and some took as little as twenty minutes.

While children were engaged in the problems, their behavior was recorded on ½-inch video tape (black and white). This is the only record that was kept. The behavior of all children in the sample produced 54 hours of taped data.

Using the Video Tape Recorder as a Data Gathering Device

There is one very obvious advantage to taped data. They represent a permanent record of the actual behaviors of the children. One doesn't run into the same problem of reliability in the interpretation of the data as is usual with other methods of observation. If there is a question, one can always return to the actual event simply by running the tape through again. However, it is time consuming to search through the mass of original data to check out some short sequence on tape. Regardless of the thought given to organizing and storing the taped information, retrieval is never easy and efficient.

At some point the information has to be taken from the tape and coded in a systematic way. The coding system can be worked out by observing the actual tapes and does not have to be done beforehand as would be the case if the data were coded while the behavior was going on. Those who are coding taped behaviors do not have to be so highly efficient in using the code from the beginning. Their skill and efficiency in using a code can be developed as they record. Reliability checks can be very quickly run and blinks can be easily applied. If a coder cannot learn to record the behaviors in a reliable way, he can be replaced with no serious harm to the research.

Another advantage to taped data is that one is readily able to detect flaws in consistency in the use of protocols. Slight variations in protocol may occur unnoticed if behaviors are being coded as they occur.

One has also to consider shortcomings of using taped data. The behavior of children when a camera is trained on them may not always be the same as it would be without the camera. Also there is a necessity to introduce extra people and apparatus into the situation which may distract the child. Fortunately in the research reported here, the problems were attractive enough to most children that such distractions were minimal. Another difficulty is that the behaviors have to occur in a relatively small space. Behavior cannot be faithfully filmed if the child is permitted to move about a great deal. It may be that for some problems, such movement would be desirable.

6 6 0 0	V ₃ "Yes"	R ^② _T R ^② _{↓2} L ₃ R ₃ L→R ^① _↓ Lp(ea. bl.) _{c.2} R ₃ L ₃ →
4 4 4 0		
4 4 4 0	E B _T	L ₆ → ^③ ₆ ₅ → ^③ ₄ → ^③ L ₃ → ^② L ₂ R ₆ L ₅ L ₄ L ₃ → ^① R ₃
2 2 2 2 2 2	R ₂ R ₁ E ₃	
	E R ₅ → ^⑥ R _↓ ^⑥ L _↓ ^⑤ V ₃ "No" E "What happens?" V ₃ "I gotta	
2 2 2 2 2 0 (holds 2) have 3 more."	Bm _{1,5} →	
3 2 2 2 3		

Note: This example illustrates the format of the code. Detailed interpretation of the code will be included in the final report to Canada Council.

It has been necessary to devise a coding scheme for each problem individually. The code must be easily read and must include those behaviors which are important but must eliminate other behaviors that are not stimulated by the problem. Coding for some problems has been completed. Below is an incomplete sample of coded behavior of a seven year old child working with the factor platform (described earlier).

Interpretation of these codes has not proved too difficult and patterns of behavior can often be detected just by looking at the coded material. If ever a particular coding is found to provide an inadequate picture of what actually occurred, a revised system can be developed without loss of data. If on-the-spot coding is done and the code doesn't work properly, all the data recorded under the defective system would be lost.

Some Further Observations

Although analysis of the data from this project is far from complete, there are a few observations which can be made at this time.

1. With the use of the guidelines developed here, it is possible to create problems which stimulate a great deal of problem solving activity on the part of young children. Even though the mathematical process involved may not have been fully understood by these children, the activity they engaged in with the material was almost always directed to some kind of solution. Only one extremely shy child in the entire sample of ninety children could not be induced to react in the problem situations. The others tended to become so involved in problem solutions that they completely ignored the rather overwhelming laboratory set-up.

An encouraging aspect of the guidelines is that once they are familiar to a person with a reasonable mathematics background and an understanding of child development, there seems to be no real limit to the number of "good" problems that, given time, the person could create.

2. Even very young children display a surprising range of activities which have mathematical significance. Making groups of objects with the same number in each group, fitting corner to corner and edge to edge to make three dimensional shapes, displaying rather complex sequences, moving about confidently in a plane, copying complex geometrical solids are only a few of the activities readily engaged in by children in the whole age range.

3. Some seemingly simple procedures which are assumed to develop early fall far short of expectations. For example, only a few children at the upper age levels showed that they had mastered the partitioning process. (From a given set of objects, putting the same number of objects in each of a given number of subsets.) Also controlling the reflection of light on an object in a plane and predicting or explaining what happens when an object or a beam of light is reflected in a plane is apparently very difficult for young children.

Children can describe the properties of a set of discrete objects if they are permitted to look at all the elements in the set at once. However, if the information is given to them an element at a time in a particular order and repeatedly, many children are not able to induce and describe any common properties of the elements.

4. Many of the older children in the sample could have solved many of the problems by making a simple mathematical computation, but in these situations, hardly any children did so. They seemed to prefer solving the problem by making moves or transformations with or in the material provided in the problem. It seemed as if they viewed the physical solution as something completely independent of any mathematical computation. For example, children who knew that 19 has no divisors would keep trying to arrange a set of 19 blocks in a number of equal piles. Or children who knew that $15 \div 3 = 5$ would use a ferry boat to take three toy cars at a time (from a set of 15 toy cars) across a river and count the number of trips. The actual moves seemed to provide the kind of verification they needed.

J. MARTIN

The University of Alberta

Controlled vs. Natural Settings—Some Implications for Behavioral Analysis and Change in Classroom Situations

This paper considers five empirical and theoretical issues which have arisen from a comparison of behavioral analyses in controlled laboratory settings and similar analyses in naturalistic human environments. The empirical bases upon which some standard behavioral variables and processes rest are found to differ slightly from one situation to the other. Implications for the behavioral analysis of classroom environments are discussed. (Dr. Martin is a Lecturer in Educational Psychology, The University of Alberta.)

A prodigious amount of literature under the general rubrics of "reinforcement theory and its applications" or "behavioral modification" has been accumulated over the past two decades. More recently, the empirical foundation supplied by such work has been incorporated by Bugelski (1964), Becker, Engelmann and Thomas (1971) and others, and logically extended to the area of classroom learning and instruction. On the basis of the research to date, it is clear that the behavioral concepts and principles developed thus far can be applied to the teaching situation and can produce gratifying results. Statements of demonstrated relationships between environmental events and behavior such as "The longer the interval between an operant response and its reinforcing stimulus, the less the strengthening effect on responses of the same class" are of obvious importance to the classroom teacher. These statements are the *facts* of the science of psychology as generated by an experimental analysis of behavior. However, while the growth of a psychology of learning and instruction built upon the results of behavioral analyses in controlled, manipulated environments may legitimately be viewed as an important contribution to educational practice, the essential extrapolation of procedures and findings to natural classroom environments can be shown to involve certain empirical and theoretical difficulties.

During the past three years, the writer has been engaged in developing methodological procedures which permit a systematic analysis of

behavioral contingencies as they operate in non-manipulated small group situations—predominately student learning groups, and psychotherapy groups (*cf.* McLeish & Martin, 1975). Taken together, these methodological innovations make possible an exacting analysis of naturalistic behavioral data, in much the same manner to which behaviors are studied in controlled environments. A comparison of the results obtained in these two settings highlights some problems which must be resolved if behavioral science is to achieve the practical success in education which its evolution as a positive science would promise. Without being “pejorative” of present practices in the psychology of learning and instruction, the paper aims at alerting educators, psychologists, and educational researchers to difficulties which urgently demand further thought and analysis. Five such issues will be reviewed.

1. *Determining the functional properties of environmental stimuli.*

For the functional behaviorist meaning is synonymous with function. The meaning of any piece of activity is determined by the sets of environmental stimuli which precede and follow upon its emission. When a baseball player swings at a pitch, he does so because he has been reinforced in the past by an ensuing drive into deep center or some such place. If a subsequent environmental event increases the frequency with which a response is emitted, it acts as a reinforcing stimulus. Its reinforcing properties are disclosed by a study of the future emission pattern of the behavior upon which it follows. This is the only valid criterion for determining whether a stimulus is reinforcing in any particular situation.

Unfortunately, in the actual practice of behavior management in the classroom (*e.g.* Zimmerman & Zimmerman, 1962; Ward & Baker, 1968; O’Leary *et al.*, 1969) this criterion is often superceded by judgments based upon the “essential” positive or negative natures of particular stimuli. On a structural basis, certain stimuli such as “warm smiles,” “friendly nods,” praise, round wooden tokens, and merit points are assumed to be positive reinforcers for all individuals at all times.

Now while there may be some consensual or statistical validity for the view that certain stimuli are more positive in more situations than are other stimuli, this is an extremely dangerous assumption. The mere fact that most of those who are subjected to behavior management programs have not managed to acquire conventional social responses to standard stimuli in their environments, should warn us that functional relationships other than conventional ones may be operating in many of these cases.

For example, in an experiment with psychotherapy outpatients as subjects (McLeish & Martin, 1975), it was discovered that aggressive interruptive behaviors were positively reinforcing for the general vocal behavior of two female subjects, whereas emphatic therapeutic responses produced behavioral frequencies indicative of the operation of punishment. Because of this and much other similar evidence, it would seem that there is no stimulus which serves a universal function, and that behavior management programs which employ such concepts, either implicitly or explicitly are doomed to failure in the specific case. No classroom program will be optimally successful unless the stimuli which are manipulated are appropriate for individual students.

2. Changes in the functional properties of stimuli over time.

The second issue is in some ways a corollary of the first. In some classroom management programs the implicit assumption is made that the functional properties of discriminative and reinforcing stimuli remain static over the temporal course of the modification program. If a particular stimulus acts as a positive reinforcer for a particular response at one point (A) in time, this does not necessarily imply that the same operant relationship will remain intact at a future point (B). Further, the time interval between points A and B need not be long. Our empirical results show that as little as five minutes (and sometimes less) may be sufficient to shape-up such a functional change.

Perhaps the best example of this phenomenon is taken from an analysis of the contingencies of reinforcement which operated during a student learning group composed of freshman students at the University of Alberta (McLeish & Martin, 1975). In the first two minutes of the seminar, the instructor's "emphatic" cephalic nods attached themselves as positive reinforcing stimuli to particular verbal responses of student A. However, this ideal theoretical pattern was eclipsed in the succeeding four minutes of interaction so that as the instructor's nods increased dramatically in their frequency of emission, they became functionally aversive, either attaching to student A's responses as punishing stimuli, or to the responses of other students as negative reinforcers in contingencies of avoidance.

Any naturalistic situation is characterized by a dynamic chaining of behavioral patterns. In the classroom, a response in one contingency for one student may act as a reinforcer for a different response by a second student and may also be caught up in a discriminative stimulus complex for a series of future responses. The competitive inhibitions and associations which result from this dynamic vortex seem to shape and modify specific stimulus properties even within relatively short time intervals.

3. Stimulus deprivation vs. extinction

Perhaps the most obvious difficulty that is encountered when behavioral analyses and techniques are extrapolated from a laboratory situation to a naturalistic human environment involves the inadequacy of traditionally established behavioral processes such as reinforcement, punishment and extinction in explaining all fluctuations in the emission rates of "social" behaviors.

McLeish and Martin (1975) found it necessary to consider a process which they have called "stimulus deprivation." Although this is certainly not an original term, its use in the present context is unique and specific.

If it were possible to hold a human social environment constant in the same manner as laboratory environments are regulated, the process of extinction would be sufficient to explain all systematic decreases in emission rates. However, because any natural social situation is dynamic in its stimulating qualities, operant responses may decrease in frequency *not* because they cease to be adequately reinforced, but because they are not subjected to the discriminative stimulus situations which come to exert control over their emission. Each time an individual in a classroom speaks or moves, his action alters the social environment in some way for other par-

ticipating individuals, and indeed for himself as well. Potential discriminative stimulus situations are thus constantly changing.

Even though a particular student is consistently reinforced for his speaking behaviors, the classroom situation is not always propitious for his contributions, and periodic decrements in the rate of emission of his speaking behavior will be witnessed. No withholding of reinforcement has occurred. What has occurred is that preceding discriminative stimuli have been withdrawn. The withdrawal of a discriminative stimulus is thus a common process in the human environment, and as such is labelled stimulus deprivation.

As a new era of sophistication in behavior technology and programming unfolds, it may be necessary to consider additional processes which have not been historically established by the experimental works of Skinner and company.

4. "Withdrawing" or "withholding"?

A fourth difficulty is that even some of the standard behavioral processes may require further clarification if they are to have optimal explanatory power in classroom situations. The concepts of punishment and extinction can be particularly troublesome.

Skinner (1953) defines two kinds of punishment—the *withdrawal* of a positive stimulus (type 1) and the presentation of an aversive stimulus (type 2). Extinction involves the *withholding* of a reinforcing stimulus. The distinction between extinction and punishment of the first type is reflected in the difference between *withholding* and *withdrawing*. When a rat in a Skinner box presses the bar and the magazine fails to click and a food pellet (the positive reinforcing stimulus) fails to drop, we say that positive reinforcement has been *withheld*. Our use of the term withheld is semantically justified by the fact that the rat, prior to this occurrence, had received a food pellet each time he depressed the bar. When a child is playing with a toy and we take the toy away from him, we *withdraw* the positive reinforcing stimulus. The difference in terminology is justified by the fact that in the case of the rat, the *withheld* food pellet is not directly perceived by the behaving organism, whereas, in the case of the child, the toy is actually in his possession and directly perceived by the senses before being removed. The withholding of the pellet in the first case leads to extinction. The number of "withholdings" is sufficient to overcome the response pattern established by whatever schedule of reinforcement was previously in effect. The withdrawing of the toy in the second case counts as punishment—type 1.

The two examples thus considered present no theoretical difficulty. However it is possible to provide numerous examples from classroom situations which defy this logical exclusivity. When a student's behavior is being positively reinforced by a smile from the instructor, and the instructor ceases to smile, is this punishment or extinction? If the smile is "removed" precisely between two successive emissions of the student's response, it would seem to count as extinction. This is because the smile is withheld after a response which has been previously reinforced. However, in dealing with a time interval which may be as small as one second, the distinction based on such arguments seems rather forced.

The student's contact with the positive reinforcing stimulus (the instructor's smile) is somewhat different from the rat's contact with the food pellet. The individual can actually observe the change in the instructor's expression as his smile disappears. The rat however, does not observe the mechanical alterations which affect the functioning of the food magazine. If the instructor emits a continuing smile during several emissions of the student's response and then alters his countenance for a subsequent emission, is this withdrawal or withholding? Clearly, more attention to this problem is warranted.

Extinction differs from punishment of the second type in a more obvious manner. In extinction a negative reinforcing stimulus may be withheld in order to curtail a contingency of negative reinforcement, but at no time does the presentation of an aversive stimulus count as extinction.

	Punishment Type I	Punishment Type II	Extinction	Stimulus Deprivation
Definition	withdrawal of +Re	presentation of aversive stimulus	withholding of +Re or -Re	withdrawal of discriminative stimulus

Another source of uncertainty is the chronological order of stimulus, response and reinforcer. If punishment is the withdrawal of a positive stimulus, stimulus deprivation would appear to be a form of punishment. Skinner, however, illustrated punishment only with examples of the withdrawal of, or presentation of, stimuli which succeed the response.

5. *Operant vs. classical conditioning*

Still another theoretical issue is concerned with the importance of preceding stimuli in controlling behavior. Skinner has established a science of behavior which is primarily concerned with the control which subsequent events exercise over responses. The effect of this is that rigid demarcation lines have been drawn between (so-called) "operant" conditioning and "classical," Pavlovian or "respondent" conditioning. This distinction has been unfortunate. It has served to limit unduly the orientation of research workers. Experiments in operant conditioning are considered by many to be valid only if they adhere to the demarcation lines. Skinner himself, while recognizing that respondent conditioning does operate in certain instances, and that even in "operant" conditioning discriminative stimuli exercise some control over behavior, tends to bypass this problem. He is more interested in carrying out an in-depth analysis of reinforcers and their effects on behavior. Nevertheless, the decision to do so seems to be arbitrary. If two events occur in a chain of other events it may be rather limiting to consider these events in terms that one event precedes the other, while ignoring the total chain of causation in which they are embedded. Skinner looks only at *half* of the conditioning process: the other half is supplied by Pavlov.

In truth, Skinner's concern with reinforcing stimuli to the exclusion of eliciting stimuli seems more methodologically pragmatic than theoretical. While it is difficult to quantify all the physical properties of an experimen-

tal space, especially in terms of temporal distances between stimuli and responses, it is less difficult to quantify the physical properties of an experimentally manipulated reinforcing event and the time intervals between responses and subsequent stimuli. This produces nice “clean” experimentation in the animal laboratory. But the clarity and cleanness are illusory. In naturalistic human settings, in the absence of experimental manipulations of variables, the research worker finds that reinforcing stimuli are *just* as difficult to recognize as are preceding (discriminative) stimuli. The pragmatic advantage of concentrating solely on subsequent events in terms of controlling behavior is lost.

There seems, therefore, to be no valid reason for limiting the scope of naturalistic investigation of classroom behaviors to *operant* phenomena, as defined by Skinner. Indeed, an overall picture of functional connections between behaviors, and between behaviors and the environment, would be enriched by a closer scrutiny of the physical aspects of the stimulus situation to which subjects are responding. The rigorous scientific work of Skinner, supplemented by the constructs of Ivan Pavlov and his model of the intact behaving organism, offers a sound base on which future work might rest.

Concluding Statement

The training of teachers in the use of behavioral procedures offers great promise for the efficient management of student learning behavior. However, behavioral scientists, in the face of severe (and all too often “emotionally-laden”) criticism from those oriented towards the concepts of phenomenology, humanism, existentialism and psycho-analysis are often prone to overstate their case, and to go beyond supportive evidences. As Skinner (1972) states:

... we are on the verge of a new pedagogy—in which the teacher will emerge as a skilled behavioral engineer. He will be able to analyze the contingencies which arise in his classes, and design and set up improved versions. He will know what is to be done and will have the satisfaction of knowing that he has done it (p. 233).

Such a pedagogy is indeed a possibility of the future, but many difficulties in the present state of behavioral knowledge must be resolved before the veracity of Skinner’s “apocalyptic vision” can be reasonably assessed.

Behavioral Analysis and Change in Classroom Situations

References

- Becker, W. D., Engleman, S., & Thomas, D. R. *Teaching: A course in applied psychology*. Chicago: Science Research Associates, Inc., 1971.
- Bugelski, B. R. *The psychology of learning applied to teaching*. New York: The Bobbs-Merrill Company, Inc. 1964.
- McLeish, J., & Martin, J. F. Verbal behavior: An experimental analysis. *Journal of General Psychology*. 1975, in press.
- O'Leary, K. D., Becker, W. D., Evans, M. B., & Saudargas, R. A. A token reinforcement program in a public school: A replication and systematic analysis. *Journal of Applied Behavior Analysis*, 1969, 2, 3-13.
- Skinner, B. F. *Science and human behavior*. New York: The MacMillan Company, 1953.
- Skinner, B. F. *Cumulative record*. (3rd ed.) New York: Appleton-Century-Crofts, Inc., 1972.
- Ward, M. H., & Baker, B. L. Reinforcement therapy in the classroom. *Journal of Applied Behavior Analysis*, 1968, 1, 323-328.
- Zimmerman, E. H., & Zimmerman, J. The alteration of behavior in a special classroom situation. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 59-60.

S . V E R M A

and

D . L . P E T E R S

The Pennsylvania State University

Day Care Teacher Practices and Beliefs

A self-administered rating scale and an observational schedule were developed to assess the beliefs and behavior of early childhood teachers in relation to constructs derived from Piagetian cognitive developmental theory and operant learning theory. Application was made to existent Piagetian and operant model early education programs. The results indicated both measures satisfactorily differentiated among the programs in a manner consistent with theory. The instruments were then applied in 38 day care centers. The results indicate that although consistent and meaningful differences are found among teachers, a wide range of internal inconsistencies exist between a teacher's stated beliefs and her actual practices. (Dr. Verma is with the Department of Education, Capitol Campus, The Pennsylvania State University at Middletown; Dr. Peters is Associate Professor of Human Development, Division of Individual and Family Studies, The Pennsylvania State University at University Park.)

During the last ten years great efforts have been made to increase the quantity and quality of early education programs. As part of this effort day care has become a topic of major concern and the number of day care facilities has increased markedly. Yet many questions concerning the process or transactions occurring within day care centers remain (Grotberg, Chapman, & Lazar, 1971). Our lack of information is the more remarkable because the analysis of process variables generally is considered an integral part of evaluation (Stake, 1970; Provus, 1969; Scriven, 1967) and quantitative descriptive studies are considered an important first step in understanding the teaching process (Rosenshine & Furst, 1973).

In order to adequately evaluate day care programs it is important to develop measurement procedures which differentiate existing programs from one another in terms of both the observable transactions that occur and the underlying assumptions which constitute the rationale for the

program. Measures of this type would permit the isolation of variables which have potential differential effects on the children enrolled. The purpose of this study was the development of appropriate and theoretically relevant measures for the naturalistic observation of teacher/child interaction patterns within day care settings and for assessing the beliefs or attitudes held by the observed teachers.

A recent extensive review of the literature uncovered only twenty-six data based observational studies of preschool teacher behavior published during the period 1963-1973 (Gordon & Jester, 1973). Of those, only one study by Prescott, Jones and Kritchevsky (1967) directly addressed interaction patterns within day care settings. Several recent studies have added to the day care list (Jambor, 1973; Green, Hollick, Knowles, Vanderkar & Winter, 1973). Yet, none of the measures reviewed fully meets the criteria we believe should be imposed when transaction measures are used within an evaluation framework (Stake, 1967). In addition to the usual psychometric requirements, an appropriate system should:

1. have a basis in some rationale or theory appropriate to early education (*i.e.* there should be some *a priori* reason for thinking the observed behaviors are important to early childhood programs);
2. provide a means for determining the internal consistency of behaviors with each other and with the program rationale (Stake, 1967);
3. be capable of distinguishing among programs in expected ways; and
4. provide an empirical basis for demonstrating that the variables selected for observation are related to child outcomes.

A brief review of prior observational systems will help to make the point. Prescott *et al.* (1967) developed a category system based on the earlier observations of Reichenberg-Hackett (1962). The resulting observation procedure included such behavioral categories as teacher direction, guidance, restriction, development of verbal skills, neutral activities and non-communications. Each category was elaborated with a system of sub-categories. While useful for describing some aspects of the day care environment as it existed in the fifty centers observed, the procedure offers only limited data for assessing the appropriateness and effectiveness of the teacher behavior observed.

Jambor (1973) extended and modified the Prescott and Jones categories to provide a more specific rationale for judging the appropriateness and consistency of teacher behavior within day care. He reorganized the behavioral categories into three broad role model headings suggested by Katz (1970): the maternal role, the therapeutic role and the instructional role. Both Katz and Jambor argue that behavior consistent with each of these roles is required to fully meet the needs of the day care child. Jambor then went on to observe the behavior of 10 nursery school and 10 day care teachers to determine the dispersion of their behavior across the three roles. While he uncovered large individual differences, he found no significant group differences between the day care and nursery school teacher groups. Both engaged in many more instructional and maternal interactions than therapeutic ones. Little relationship was found between any of the teaching roles and a variety of classroom and teacher characteristics.

Though based on practical logic and providing a means for judging the

internal consistency of behaviors, Jambor's categories lack both a theoretical foundation and prescriptive power as they have been developed to date. The data produced are not immediately suggestive either for day care program improvement or for revision of training programs for day care personnel. No attempt was made to relate any of the behaviors to child outcomes.

The *Teacher Practices Observation Record* (TPOR) developed by Brown (1968) was also designed to assess the relation of teacher behaviors to a standard, in this case Dewey's experimentalist position. The TPOR consists of 62 items, half of which represent the experimentalist position and half do not. Soar (Soar & Soar, 1972), using a modification of the TPOR for comparison among Follow-through programs, factor analyzed his data and identified several factors which identify teacher behavior with reference to program type (*e.g.* Bereiter-Englemann *vs.* British Infant School).

Although one can question the relevance of Dewey's experimentalism to the day care setting, Soar's results do suggest that programs may be readily differentiated, and that training a teacher in a particular program model tends to lead to more uniform behavior. The procedure does not, however, indicate if the uniformity is consistent with the specific rationale of the program. Soar and Soar (1972) do find that several behavior factors relate to pupil achievement.

An alternative direction is suggested by the research of Harvey, White, Prather and Hoffmeister (1966). These researchers were interested in assessing the relationship between a teacher's own beliefs and her classroom behavior. Using a sentence completion procedure the beliefs of 168 preschool teachers were assessed. Thirty teachers representing three different belief groups (high abstract, low abstract and high concrete) were then observed using a twenty-six dimension rating scale. The rating scale included such dimensions as anxiety, punitiveness, warmth, flexibility and so forth. The comparison of the three belief groups' ratings indicated there was a significant relationship between beliefs and the ratings of behavior. However, there is no attempt to relate the teacher's beliefs to any generic early childhood rationale, nor is there an attempt to relate either beliefs or behavior to child outcomes.

The studies reported here constitute a preliminary attempt to further identify variables which have potential for discriminating among early childhood programs and to develop transactional measures which meet the criteria we have imposed.

Development of Measures

Theoretical Framework

The preceding literature review suggested that an appropriate point of attack for assessing the sources of differences across programs would be teacher behaviors *and* the belief system upon which they are based. That is, since program rationales differ, if one is to assess congruence of teacher behavior with the program rationale rather than some external standard, then the rationale (beliefs) need to be measured as well as the behavior. Further, if the variables isolated for study are to have potential impact on the children enrolled, it was felt that they should be derived from the

applications of developmental theory: theory relevant for early childhood program development.

The theories of Piaget and Skinner are vastly different but both have contributed in a major way to the development of early childhood programs during the past ten years. Both have been cited extensively in the early childhood literature and have been given a place in the rationale for most, if not all, recent early childhood program developments. For this reason they were selected as the basis for instrument development.

Piagetian and Operant theory formulate different assumptions about children, development and learning (DeVries, 1974). A careful review of the literature related to each suggested ten dimensions of potential value in that the applications derived from the two theories were quite different (Verma, 1973) (See Table 1), and belief statements and observable behavior categories could be generated from each. Five of the dimensions were selected for study based on an estimate of their importance to the theory and the practicality of their use. A number of belief statements and behavior categories were generated for each dimension using items from Stern and Gordon (1967) and from Brown (1968) whenever possible.

TABLE 1
THEORETICAL ASSUMPTIONS OF TWO THEORIES

	Piaget's Theory	Operant Theory
Children	*1. Active	1. Passive
	*2. Qualitatively unlike adults	2. Qualitatively like adults (adult miniatures)
Development	3. Qualitative	3. Quantitative
	4. Interaction between child and environment	4. External environment
Learning	*5. Intrinsic motivation	5. Extrinsic motivation
	6. Based on sensory education	6. Based on language overt and covert verbal labeling
	7. Stage dependent	7. Knowledge and familiarity with task or tasks similar to it
	*8. Based on massive general type of experience	8. Based on specific training
	*9. Process approach	9. Product approach
	10. Irreversible (invariant)	10. Reversible

* Dimensions retained in final instruments.

Teacher Belief Rating Scale

A preliminary version of the Teacher Belief Rating Scale (TBRS) was developed by converting the 6 items generated from each of the five selected dimensions to a Likert-type scale. Piagetian and Operant items were randomly ordered and rating order was reversed intermittently to avoid the formulation of response sets. Two scores per respondent were attained by summing Operant and Piagetian items separately (with appropriate reversals).

The scales were then submitted to the program consultant or subject matter specialist, the head teacher and the assisting teacher of the Piagetian and Operant based laboratory early education programs of the University. As such these two groups ($N=3$ per group) were known to differ on the basic constructs being measured. Analysis of variance revealed the effects of scale to be significant ($F = 20.5$; df 1, 8; $p < .05$). The group by scale interaction was also significant ($F = 36.8$; df 1,8; $p < .05$). This interaction was considered particularly crucial for determining whether the TBRS was performing properly.

Suggestions of the experts and of eight day care teachers concerning readability and applicability resulted in the deletion of six items. The final version of the TBRS therefore consisted of 24 items (12 representing operant beliefs and 12 representing Piagetian beliefs). This version was subsequently given to eleven early childhood personnel known to be affiliated with Operant, Piagetian or other programs. Analysis again indicated as scale main effect ($F=19.8$; df 1,8; $p < .05$) and a scale by group interaction ($F = 34.5$; df 3,8; $p < .05$).

Teacher Practices Observation Form

The Teacher Practices Observation Form (TPOF) was developed by formulating observable behavioral categories each of which correspond to the items on the belief scale. As such, the TPOF underwent a parallel development to the belief scale. The initial observation form consisted of 30 items (15 Piagetian and 15 Operant—six from each of the five dimensions). After successive trials in three early childhood education programs, six categories were deleted and the remaining 24 were collapsed into 12 numbered items with a Piagetian, an operant, and a not observed alternative under each. Thus the observer was required to go down the list, checking one of the alternatives for each item immediately following a 15 second observation segment.

The TPOF then was used to collect data from the Piagetian and Operant laboratory early education programs using the procedures indicated above. Since the programs observed were training programs, observations were collected of the behavior of both the teachers and the student teachers. Average scores were calculated by dividing the total checks per item in each program with the total time observed. This procedure provided a rough rate measure for each of the teacher behaviors. Directionality was hypothesized on the basis of the scale from which a particular item was drawn. A sign test was used to determine if the relationship between rates of observed practices were consistent with the directions hypothesized by theory. Twenty of the 24 items showed differences in the predicted directions. Two items were not observed in either program and two items showed a reversal of the predicted direction. In general, the results indicated

that the observation form was successful in differentiating the Piagetian teachers' behavior from Operant teachers' behavior even when some of the teachers are not yet fully trained.

In another study (Cohen, Peters & Willis, 1974), the behaviors of student teachers were observed in each of four early education programs using the TPOF. A total of 120 observations in each room was made over a two month period. The observed behaviors for each scale were summed across student teachers by program. The total observed frequency of Piagetian and Operant behaviors are presented in Table 2. These data tend to further confirm the validity of the TPOF and its ability to differentiate across early education programs.

TABLE 2
OBSERVED FREQUENCY OF PIAGET AND OPERANT BEHAVIOR
FOUR PROGRAMS

Behavior	Program Type			
	Operant	Piagetian	Responsive Environment	Day Care
Piaget Behavior	64	472	433	361
Operant Behavior	619	96	161	206

Reliability of the Measures

Reliability of TBRS. Kuder-Richardson Formula 20 internal consistency reliabilities were estimated for the TBRS utilizing data from the total study sample ($N=38$). Two methods were used. In method one each agreement statement (slightly agree, moderately agree, strongly agree) were coded as "2" for each scale separately. Disagreement statements were coded as "1". The reliabilities thus calculated were .44 and .58 for the Piaget and Operant scales respectively. The intercorrelation of the two scales was .13 (non-significant).

In the second method Kuder-Richardson Formula 21 internal consistency reliabilities were calculated using the full 6 point scale range. In this case the reliabilities were .56 and .66 for the Piaget and Operant scales respectively. The intercorrelation of the two scales was -.04 (non-significant).

Inter-observer agreement for TPOF. 68 observations were used to pretest the final Teacher Practice Observation Form. Two observers simultaneously observed the same teacher 68 times (investigator made 68 observations and observer A and B made 42 and 26 observations respectively). The percentage of agreement between the observers was calculated.

The mean agreement across observers ranged from .72 to .96 and across items from .60 to 1.00. The overall mean agreement was .83.

Application to Day Care

Sample

The sample for the study consisted of 38 presently employed teachers in day care centers in Pennsylvania. The centers were selected from existing

lists using the criteria that the programs should be geographically representative and have more than 15 children currently enrolled.

Procedures

Two trained observers were used in the study, each visiting approximately half the programs. During the visit the head teacher was observed between 9:30 a.m. and 11:30 a.m. The observations were made during three half-hour periods with a ten minute break between them. Twenty observations were made during each half hour period using a point-sampling procedure. This yielded a total of sixty observations for each teacher. Throughout the visit the observers remained as unobtrusive as possible, though the teachers naturally knew they were being observed. After the observations were completed the teachers were requested to respond to the Teacher Belief Rating Scale.

Results

The results of the field study indicate that the scores on the belief scale and the frequencies of observed behavior were normally distributed across the sample. No clear Piagetian groups or operant groups of teachers could be discerned on the basis of either beliefs or practices. Spearman rank correlations were calculated for the total scores on Piagetian beliefs and practices, and for total operant beliefs and practices. These data indicate that only a weak relationship exists. Teachers who rank high in Piagetian beliefs also tend to rank high in Piagetian practices ($r=.30$, $p < .05$). Although a positive relationship is indicated for the correspondence of ranks on operant beliefs and practices the results do not reach acceptable levels of significance ($r=.10$, $p > .05$). A significant negative correlation was found between the ranking of Piagetian practices and Operant practices as might be expected from the nature of the construction of the observation instrument ($r=-.36$, $p < .05$).

In general the results indicate that the day care teachers agree significantly more with Piagetian beliefs ($\bar{X}=51.9$, $sd=8.57$) than with operant beliefs ($\bar{X}=42.1$, $sd=10.0$) ($t=4.69$, $p < .05$), but behave in ways more consistent with operant theory ($\bar{X}=113.3$, $sd=28.9$) than with Piagetian theory ($\bar{X}=71.7$, $sd=35.7$) ($t=5.77$, $p < .05$).

Both these findings suggest that there is little correspondence between teachers' professed beliefs and observed practices.

To adequately assess the internal consistency of each teacher's beliefs and practices simple correlation coefficients of Piagetian beliefs and practices and Operant beliefs and practices were computed for each teacher. The results indicated that only two of the thirty-eight teachers had internally consistent beliefs and practices: one Piagetian and one Operant.

Discussion

It is not surprising that no clear Piagetian or Operant groups of teachers were found in operating day care programs. Nor is it surprising that only a weak relationship between behavior and beliefs was uncovered. Such findings have been reported before (Soar & Soar, 1972; Brown, 1968). Further, little effort has been made to implement theoretically pure or internally consistent programs in day care. Rather each teacher has attempted to put together her program in the way that appears to work best for her. When the range of alternatives available in materials and methods is con-

sidered, it is not surprising that the decisions and choices are difficult to make.

Perhaps the teachers of this sample chose the operant practices more often because they involve teaching techniques which are easier to learn and take less time to implement. Other studies have suggested this might be the case (Bissell, 1971; Miller *et al.*, 1970). Yet the choices made have led to an unnatural split between theory and practice. Only two teachers had beliefs and practices that were internally consistent. It appears that theory is little applied as either a policy for guiding practice or as an after-the-fact explanation for practice. This results in both the loss to practice of what is so promising in theory and the loss to theory of what is learned in practice.

The importance of the relationship between theory and practice has been stressed recently (DeVries, 1974; Parker, 1974; Peters, Cohen & McNichol, 1974; Willis, Cohen & Clement, 1974; EPIE, 1972). An explicit and solidly based theoretical stance is seen as an important ingredient for the improvement of early childhood programming. Theory provides a guideline for planning and acting in consistent ways. Through recourse to theory the teacher can derive program goals, set priorities, develop appropriate teaching methods, select materials, deductively arrive at solutions to everyday problems, and predict appropriate responses to future situations. Operating without theory is coping; i.e., dealing with each situation as it arises.

Both the lack of a strong relationship between theory and practice and the lack of internal consistency in either, suggest the need for an increased emphasis on theory in the training of day care personnel. Current emphasis on skill training should not be divorced from the rationale underlying the selection and use of specific skills in given circumstances. Day care personnel are in need of the "whys" as well as the "how to's" if they are to function independently in planning and implementing quality day care programs. A broad range of competencies must be considered if teachers are to be thoroughly trained (Peters & Dorman, 1974).

Conclusions

The studies reported here were designed to further identify variables which have potential for discriminating among early childhood programs and to develop transactional measures which relate theory and practice. The results of the studies conducted to date indicate that there are discernibly different implications for practice that can be derived from Piagetian theory and Operant theory. Both theories have relevance for early education and potential for the development of measures which can differentiate among program practices. From the theoretical framework provided by the theories two measures were developed—The Teacher Belief Rating Scale and the Teacher Practices Observation Form. When applied to programs known to differ on their theoretical bases, both measures produced reliable data which differentiated among the programs in meaningful ways. When applied to the day to day operations of day care, theoretical purity was found neither in beliefs nor in practices.

Our studies are not yet complete. We have not yet met our own fourth criterion. That is, the relationship between the teacher behavior variables

selected for study and appropriate child outcome variables has not yet been determined. However, the progress made to date suggests that transaction or process measures for cross-program and within program evaluation can be derived by the steps we have followed. The TBRs and TBOFs appear to have reliability, validity and utility in day care program analysis. The results obtained from the major field test are suggestive for both future research and for the training of day care personnel. However, the final decision as to their importance awaits empirical assessment of the relationship between these teacher variables and changes in children.

This report is adapted from a dissertation conducted as partial fulfillment of the requirements for the Ph.D., College of Human Development, The Pennsylvania State University. The research was partially supported by the Department of Public Welfare of the Commonwealth of Pennsylvania. The opinions expressed are those of the authors and do not necessarily express those of the sponsoring agency.

References

- Bissell, J. S. *Implementation of planned variation in head start*. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Child Development, April, 1971, ED 052 845.
- Brown, B. B. *The experimental mind in education*. New York: Harper, 1968.
- Cohen, A. S., Peters, D. L., & Willis, S. L. The effects of early childhood education student teaching on program preferences, beliefs, and behaviors. (In preparation, 1974.)
- DeVries, R. Theory in educational practice. In R. Colvin & E. Zaffiro (Eds.), *Preschool education: A handbook for the training of early childhood educators*. New York: Springer, 1974, 3-40.
- Educational Products Information Exchange. *Early Childhood Education. How to select and evaluate materials*. Educational Product Report #42 New York: EPIE, 1972.
- Gordon, I., & Jester, R. Techniques of observing teaching in early childhood and outcomes of particular procedures. In R. Travers (Ed.) *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.
- Green, R., Hollick, R., Knowles, K., Vanderkar, C., & Winter, M. The Pennsylvania day care study survey questionnaire and observation reports. Technical Report #9. Center for Human Services Development, The Pennsylvania State University, 1972.
- Grotberg, E., Chapman, J., & Lazar, J. A review of the present status and future needs in day care research. (A working paper prepared for the Interagency Panel on Early Childhood Education Research and Development) Office of Child Development, H.E.W., Washington, D.C., November, 1971.
- Harvey, O., White, B., Prather, M., & Hoffmeister, J. Teachers' belief systems and preschool atmospheres. *Journal of Educational Psychology*, 1966, 57, 373-381.
- Jambor, T. Instructional, maternal and therapeutic role behavior of day care and nursery school teachers. Unpublished doctoral dissertation, The Pennsylvania State University, 1973.
- Katz, L. Teaching in pre-schools: Roles and goals. *Children*, 1970, 17, 42-48.
- Miller, L. B. et al. *Experimental variation of head start curricula: A comparison of current approaches*. Progress report number 5. Research grant #0158189, Office of Economic Opportunity, Louisville, Ky.: University of Louisville, January, 1970 (a), ED 041 617.
- Parker, R. K. Theory in early education curricula. In R. Colvin & E. Zaffiro (Eds.) *Preschool education: A handbook for the training of early childhood educators*. New York: Springer, 1974, 41-74.

Day Care Teacher Practices and Beliefs

- Peters, D. L., & Dorman, L. Program planning and implementation. In R. Colvin & E. Zaffiro (Eds.) *Preschool education: A handbook for the training of early childhood educators*. New York: Springer, 1974, 115-137.
- Peters, D. L., Cohen, A. S., & McNichol, M. The training and certification of early childhood personnel. *Child Care Quarterly*, 1974, 3, 39-53.
- Prescott, E., Jones, E., & Kritchevsky, S. Group day care as a child-rearing environment. Pacific Oaks College, Pasadena, California, 1967.
- Provus, M. Evaluation of ongoing programs in the public school systems. In R. Tyler (Ed.) *Educational evaluation: New roles, new means*. (Sixty-eighth yearbook of the National Society for the Study of Education, Part II.) Chicago: University of Chicago Press, 1969.
- Reichenberg-Hackett, W. Practices, attitudes and values in nursery group education. *Psychological Reports*, 1962, 10, 151-172.
- Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In R. Travers (Ed.) *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.
- Scriven, M. *The methodology of evaluation: Perspectives on curriculum evaluation*. (AERA Monograph Series on Curriculum Evaluation, No. 1.) Chicago: Rand McNally, 1967.
- Soar, R., & Soar, R. An empirical analysis of selected Follow Through programs: An example of a process approach to evaluation. In I. Gordon (Ed.) *Early childhood education*. (The Seventy-first Yearbook of the National Society for the Study of Education). Chicago: University of Chicago Press, 1972.
- Stake, R. The countenance of educational evaluation. *Teachers College Record*, 1967, 68, 523-540.
- Stake, R. Objectives, priorities and other judgment data. *Review of Educational Research*, 1970, 40, 181-212.
- Stern, V., & Gordon, A. *Development of observation procedures for assessing preschool classroom environment*. Research Division Bank Street College of Education. New York, 1967.
- Verma, S. A study of the relationship between day care teachers' beliefs and teaching practices. Unpublished doctoral dissertation. The Pennsylvania State University, 1973.
- Willis, S. L., Cohen, A. S., & Clement, J. W. Formative evaluation in a cognitive developmental program for young children. Paper presented at the Fourth Annual Invitational, Interdisciplinary Conference: Piagetian theory and its implications for the helping professions, February, 1974.

ANNE D. FORESTER

University of Victoria

Learning the Language of Reading —An Exploratory Study

The general aim of the study was to gain insights into the thought processes and learning strategies of the beginning reader. It is based on the premise that since reading is a language task, the process of learning to read may show parallels to the process of learning to speak. Accordingly, the study was modelled after an observational study on the acquisition of oral language which yielded significant insights into the manner in which infants process oral language. It served as a pilot study for more extensive observations which have already confirmed some of the earlier findings.

Analysis of data collected during observations in a first-grade classroom suggest that there is a similarity between the processes of learning oral language and learning to read. The beginning reader, like the child learning to speak, appears to generate rules of his own to deal with the language input. Oral practice and feedback from the teacher guide learning more than the assimilation and application of rules presented in the classroom. Observations suggest that knowing the rules about sounds, letters, or endings follows rather than precedes the ability to manipulate these elements in the printed text. The child learns to sound out words once he has learned to read. (Mrs. Forester is a graduate student in Educational Psychology, the University of Victoria, British Columbia.)

The processes of learning to read and learning a language have a number of common elements: sounds and symbols have to be matched, new patterns of communication assimilated and tested for accuracy and semantic value. In both cases, the learner is faced with the need to assimilate and apply a multitude of rules in order to gain a degree of fluency. A child learns his native language without the benefit of formal instruction, seemingly without effort, and certainly without being apprised of the rules which govern that language. The child himself abstracts the patterns and regularities, and learns to transform them into new and sensible utterances. The beginning reader is faced with much the same task, that of learning to manipulate the rules of written communication, "making

sense" out of the language of reading. But in his case, rules and explanations are provided to help him learn.

In recent years, a number of studies have shed light on the processes involved in the acquisition of language and in view of the similarities referred to above, similar studies on the acquisition of reading appeared promising. A clearer understanding of the processes involved in learning to read might reveal to what extent the beginning reader—like the child acquiring language—is generating his own rules. If parallels would be found to exist between the two processes, such information could have a great deal of bearing on the type and number of rules provided in reading instruction and the timing at which they should be introduced.

The study here described¹ was undertaken in the hope that naturalistic observations in the classroom might provide the same type of data for analysis as that collected in the homes of infants learning to speak, in order to gain an insight into the thought- and learning-processes of the beginning reader. It was proposed as a small-scale replication of the study undertaken by Roger Brown and Ursula Bellugi (1964) on the child's acquisition of syntax, and has served as a pilot model for more extensive observations now under way with the use of VTR equipment which are corroborating and expanding some of the findings reported here.

In their study on the acquisition of syntax, Brown and Bellugi used tape recordings and direct observations of the verbal interactions between the subjects they studied and the subjects' mothers. The material thus obtained was then transcribed and analyzed to determine what patterns or processes could be inferred. Three processes were isolated: Imitation and Reduction, Imitation with Expansion, and Induction of the Latent Structures. Analysis of the latter process in particular stressed that inner processes of structuring more than imitation of adult patterns accounted for the acquisition of the more complex syntactic structures. It is described as "the very intricate simultaneous differentiation and integration that . . . is more reminiscent of the biological development of an embryo than it is of the acquisition of a condition reflex (Brown & Bellugi, 1964, p. 151.)"

The proposal to study the acquisition of reading in the same manner, the use of tape recordings of the interactions between the subjects and the teacher, proved impractical. The general classroom noise level interfered with the taping and it became necessary to limit this initial study to general observations and shorthand notes about the interactions of significance. Rather than limiting the observations to two to four subjects after initial visits to the classroom, all the children were observed, their interactions with each other being included as well.

The Classroom

The study was undertaken in the first-grade section of Fairburn School in Victoria, B.C., which was composed of twenty first-graders (12 boys, 8 girls) and, during the afternoon, eight kindergarten children (2 boys, 6 girls). The children were of the appropriate age (5 and 6 years old respectively) and came principally from middle-class homes. The atmosphere in the classroom was friendly and the children seemed eager to respond to

¹ The study was undertaken under the guidance of Dr. Norma I. Mickelson and Dr. Peter O. Evanechko of the Faculty of Education, University of Victoria.

questions. They were accustomed to outside visitors and did not appear to be worried about or overly interested in the presence of an observer.

The programme of instruction was one of integrated language arts combining phonics, sight-word instruction, reading in several texts, and writing.

Data Collection

Observations in the classroom began on February 1st, 1974, and extended to March 20th. Generally, two visits per week were made during that period; however, on several occasions other activities (French lessons, special events, field trips, teacher conferences, etc.) prevented observations of activities connected with the acquisition of reading. Ten sessions of about one-and-a-half hours each provided the material discussed here. Three of the sessions were devoted to work with kindergarten children. As noted above, a written record was kept of interactions between the teacher and pupils. This included blackboard work, oral interactions during reading sessions and exercises in connection with worksheets. These observations were augmented by notations made during oral readings in class and tape recordings of oral readings by three children selected by the teacher and designated as top, middle, and poorest readers among the first-graders. Analysis of the oral readings, both those of the three pupils during special sessions and of groups of children during regular class work, was done on the pattern of the Goodman Miscue Analysis² (Goodman, 1969).

Observed Behaviour Patterns

Analysis of the data collected—records of oral interactions and notations on general reading behaviour—revealed a number of recurring patterns which are listed below with selected examples or explanatory comments.

1. The same wrong answer, or one very similar to it, is frequently given several times in succession by a number of children or even the same child. (The answer looked for by the teacher is usually readily inferred by an adult observer, but apparently not by the children.) The behaviour described in 2 below is a case in point.
2. The child's perceived meaning and his customary names for concrete referents appear to guide his response even after the teacher has rejected an answer several times. The task at hand—say, finding rhyming words or beginning sounds on an exercise sheet—does not intrude itself on the recognition of pictures as “glove” or “beetle” when “mitten” and “bug” are the appropriate responses for that particular exercise.
3. The intended purpose of an exercise—say, learning about letters and sounds—does not necessarily guide the child in working out the answers. On a sheet showing pictures of rhyming words the picture of a cat appeared in four locations. Since “cat” was used to rhyme with “hat” in the first row, a child stated that that meant “cat” was ruled out as a

² In Goodman's view, differences between the expected response and observed response during oral reading yield important insights into the mental processes of the reader. They offer glimpses of reading “from the inside out.” Instead of classifying such differences between the text and the oral rendition of the child as “errors” they are termed “miscues” and used as an analytic tool—with the aid of a taxonomy of miscues—to derive an understanding of the child's thought processes and degree of comprehension of the material read.

rhyming word for the rest of the paper. ("Cat" was, in fact, used only once as a rhyming word in the work sheet.)

4. Statements by other children that "this is the right answer" or demonstrations that "this is the way to do it" appear to be more readily accepted and copied than explanations of an adult—teacher or observer.
Jimmy: How do you spell cauze?
Adult: Do you mean because?
Jimmy: No, CAUZE! (enunciating emphatically)
Other child: Yeah, it's "because," Jimmy.
Jimmy: Oh. O.K., how do you spell because?

5. Explanations provided by the teacher about such concepts as words and sentences frequently fail to elicit a response or sign of recognition. Children appear to deal with language in patterns and semantic units. Classification of the language into discrete grammatical units at the word level has not yet been internalized.

When asked by the teacher to give their favorite word, two of the kindergarten children responded with "happy face" and "swimming lesson." Their responses to the teacher's explanation that those were two words were respectively, "Then I don't know." and a repetition of "swimming lesson." A first-grader faced with a picture and the word "catch" read it as "Are you ready?" in response to the teacher's request to "read the word."

6. "Sounding out" a word does not yet appear to proceed in the manner conceived by the proficient reader. The children frequently appear puzzled about the connection between "the sound" the teacher is demonstrating or talking about and the word they are asked to read. They will cup their lips to imitate the teacher and even produce "the sound," nod "yes" to the question "do you hear the sound?" yet fail to make the transfer from "the sound" to the word.
7. Errors (miscues) during oral reading sessions frequently involve conversion of the text into a language pattern more meaningful in terms of the child's oral language. "Mother" becomes "Mom," "Little Duck" is read as "baby duck," and "No, I have not seen her." is shortened to "No, I haven't."
8. Wrong intonations during oral readings often indicate that the child draws on his knowledge of familiar language patterns or word order to derive meaning from the text. "Father will be back tonight. The children . . ." (In the text, the sentence appeared in the same manner—with "tonight" beginning the next line) becomes "Father will be back. Tonight, the children . . ." The sentence "Then you push a little." created difficulty for a first-grader. He stopped after "push a . . ." despite the fact that "little" is a very familiar word to him. When he was coaxed into reading the sentence his intonation indicated that to him the sentence was not complete. "Little" apparently required a complement.

Discussion

The short duration of the study and the necessity to observe an entire class with little more than the aid of a pencil and paper made it impossible to draw definitive conclusions. However, even these limited observations

confirmed that there are indeed parallels between the processes of language acquisition and learning to read. Brown and Bellugi (1964) found that some of the ungrammatical utterances produced by the infants they observed yielded particularly interesting data for analysis since the "inductive operations of the child's mind are externalized in such a creation." Similarly, reading errors and the children's overall reactions to instruction in the classroom at times externalized their thinking. It revealed that from a very early stage, beginning readers, like infants learning to speak, draw upon the totality of their environment and experience to "make sense" of language: in the classroom, children look to the teacher and to other children to confirm the correctness of their efforts with written language; they examine pictures in the text, and, above all, they draw on their knowledge of their own patterns of oral language.

Where for the infant, contentives (nouns, verbs, and adjectives) dominate early speech, and language patterns (word order, noun phrases, etc.) appear to guide the further development of language, so the beginning reader seems to rely on concrete referents and his own language patterns to deal with written language (see examples under 2, 5, 7, & 8 in the foregoing section.)

Oral practice rather than reliance on explanations or abstract concepts appears to guide the learners in most cases. The infant practises speech on the basis of the oral language he hears around him. He receives feedback on his efforts and gradually approximates the adult models' speech more closely. He himself abstracts the rules, tests them orally, (Child: Daddy goed out. Mother: That's right, Daddy went out.) and eventually gains proficiency in their use; yet the rules themselves are never explicitly stated. In class, beginning readers repeat the teacher's examples, receive feedback on their own performance and that of other children in the class. Their ability to perform correctly improves gradually even though they cannot accurately verbalize the explanation the teacher has repeated many times. During exercises involving word endings, most children performed correctly but were unable to state reasons for their correct responses. Examples collected several weeks apart revealed that proficiency in the use of apostrophe s to indicate possession had improved, yet the children continued to use a tautology to explain their choice: "There is an apostrophe s on the end because it's Ann's doll." None repeated the teacher's explanation, "The doll belongs to Ann."

During phonics lessons, children perform correctly in a high proportion of cases but display a certain rigidity of response. Often, the more proficient readers will provide the initial examples and others follow suit. However, at times, responses quite unlike those desired are given and the teacher's question, "Do you hear the sound?" brings out the fact that the connection between "the sound" and the word is not perceived by a child who can nevertheless read.

Since the child is unclear about such concepts as sounds, words, sentences, and context, it becomes apparent that he largely abstracts and integrates the rules of reading on the basis of his own cue system aided by the oral practice in class. Much like the child learning to speak, the beginning reader inductively arrives at latent structures in the language (in this

instance written language). In this process, the explanations offered by the teacher may take a role akin to that of the mother's expansion of the child's utterances during the "imitation-with-expansion" phase of the process of language acquisition. As the child is practicing his new skills in class, the teacher acts as "communication check" to confirm, expand on, or correct his efforts.

On the basis of research now under way, further analogies between the processes of language acquisition and reading acquisition are beginning to emerge. In addition to Brown and Bellugi's study, some facets of Lenneberg's (1969) biological theory of the development of language should prove useful in formulating a more rigorous theoretical basis for the analysis of data collected during classroom observations. His concepts of gradual development, the actualization of latent structures into realized structures, the function of the social environment as a triggering mechanism (rather than a causal agent), and the process of assimilation, could readily be applied to a model of reading acquisition.

Conclusion

Analysis of data collected during observations in a first-grade classroom appears to confirm that there is a similarity between the processes of learning a language and learning to read, though the input provided for learners differs. The beginning reader is guided by explanations regarding phoneme-grapheme correspondences and the rules of syntax and grammar. The infant learning to speak, acquires proficiency without the benefit of explanations or stated rules. Both have the opportunity to practice orally and both receive feedback on their efforts. The finding that despite the differences in the learning environments, beginning readers nevertheless appear to use strategies similar to those of the beginning oral language user, raises the question to what extent the child in the classroom needs and uses the rules presented to him during the initial efforts to penetrate written language.

If reading is defined as the process of deriving meaning from the printed word, then an untimely introduction of rules and abstract explanations may actually retard the process of learning to read. While the child is searching for meaning, we, the teachers, are fixating on "sounds", letters, and words. As Frank Smith (1974) put it, "If we tried to teach children to speak the way we teach them to read, they might never learn." Assuming the necessity to "know the rules" before being able to perform, seems to be putting the cart before the horse. The external input presented by the teacher serves as a model and provides that all-important feedback on the child's performance; however, it would appear that it is not that the child learns to read by sounding out words, but rather that he learns to sound out words once he has learned to read. He knows the right answers before he knows why he is right. He learns the language of reading before he learns the rules.

References

- Brown, R., & Bellugi, U. Three processes in the child's acquisition of syntax. *Harvard Educational Review*, 1964, 34, 133-151.
- Burke, C. L., & Goodman, K. S. When a child reads: A psycholinguistic analysis. *Elementary English*, 1970, 47 (1).
- Goodman, K. S. A linguistic study of cues and miscues in reading. *Elementary English*, 1965, 42.
- Goodman, K. S. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly*, 1969 (Fall).
- Goodman, Y., & Burke, C. Do they read what they speak? *Grade Teacher*, 1969, 26(7).
- Lenneberg, E. H. *Biological foundations of language*. New York: John Wiley and Sons, Inc., 1967.
- Smith, E. B., Goodman, K. S., & Meredith, R. *Language and thinking in the elementary school*. New York: Holt, Rinehart & Winston, 1970.
- Smith, F. *Understanding reading*. New York: Holt Rinehart & Winston, 1971.
- Smith, F. *Psycholinguistics and reading*. New York: Holt Rinehart & Winston, 1973.
- Smith, F. Address given at the University of Victoria, Victoria, Canada, Spring, 1974.

Editor's Note:

It has not been a general policy of AJER to publish bibliographies. However, this one, developed over the last several years by Dr. Blackall, represents an area which we think will be of wide interest to our readers.

BETTINA M. BLACKALL

The University of Alberta

The Works of Jean Piaget Published in English

Piaget, J. *The language and thought of the child.*

Translated by Marjorie Gabain. London: Routledge & Kegan Paul, 1959. Original French edition, 1923; First published in English, 1926.

London: Routledge & Kegan Paul, 1926, 1948.

New York: Harcourt, Brace, 1928.

London: Routledge & Kegan Paul, 1952.

New York: World Books, 1955.

New York: Meridian, 1957.

London: Routledge & Kegan Paul, 1959. (pb)

New York: Philosophical Library, 1961.

New York: Humanities Press, 1962. (Reproduction of the 1929 ed.)

Piaget, J. *Judgment and reasoning in the child.*

Translated by Marjorie Worden. New York: Littlefield 1968. Original French edition, 1924; First published in English, 1928.

London: Routledge & Kegan Paul, 1928.

New York: Harcourt, Brace, 1928, 1938.

London: Routledge & Kegan Paul, 1951.

New York: Humanities Press, 1962. (Reproduction of the 1928 ed.)

New York: Littlefield, 1968, (pb)

Piaget, J. *The child's conception of the world.*

Translated by Joan & Andrew Tomlinson. New York: Littlefield, 1969. Original French edition, 1926; First published in English, 1929.

London: Routledge & Kegan Paul, 1929.

New York: Harcourt, Brace, 1929.

London: Routledge & Kegan Paul, 1951.

New York: Humanities Press, 1960. (Reproduction of the 1929 ed.)

Totowa, N. J.: Littlefield, Adams, Patterson, 1960.

New York: Littlefield, 1969. (pb)

London: Paladin, 1973. (pb)

Piaget, J. *The child's conception of physical causality*.

Translated by Majorie Gabain. New York: Littlefield, 1965. Original French edition, 1927; First published in English, 1930.

London: Routledge & Kegan Paul, 1930.

New York: Harcourt, Brace, 1930.

Totowa, N. J.: Littlefield, Adams, Patterson, 1960.

New York: Littlefield, 1965. (pb)

New York: Humanities Press, 1966. (Reproduction of the 1930 ed.)

Piaget, J. Children's philosophies. In C. Murchison (Ed.), *A handbook of child psychology*. Barre: Clark University Press, 1931. Pp. 337-391.

Piaget, J. Retrospective and prospective: Analysis in child psychology. *British Journal of Educational Psychology*, 1931, 1, 130-139.

Piaget, J. *The moral judgment of the child*.

Translated by Marjorie Gabain. New York: Macmillan, 1955. Original French edition, 1932; First published in English, 1932.

London: Routledge & Kegan Paul, 1932.

New York: Harcourt, Brace 1932.

New York: Macmillan, 1955.

Piaget, J. *Social evolution and the new education*. (Vol. 8) London: New Education Fellowship, 1932.

Piaget, J. Principal factors determining intellectual evolution from childhood to adult life. In *Factors determining human behavior*. Harvard Tercentenary Publication, I. Cambridge, Mass.: Harvard University Press, 1937. Pp. 32-48.

Piaget, J. Intellectual evolution. In *Science and man*. Totowa, N. J.: Littlefield, Adams, Patterson, 1942. Pp. 409-422.

Piaget, J., & Inhelder, B. Diagnosis of mental operations and theory of intelligence. *American Journal of Mental Deficiency*, 1947, 51, 401-406.

Piaget, J. *The psychology of intelligence*.

Translated by M. Piercy & D. E. Berleyne. New York: Littlefield, 1968. Original French edition, 1947; First published in English, 1950.

London: Routledge & Kegan Paul, 1950.

New York: Harcourt, Brace World, 1950.

Totowa, N. J.: Littlefield, Adams, Patterson, 1960.

New York: Littlefield, 1968. (pb)

Piaget, J., & Weil, A. M. The development in children of the idea of the homeland and of relations with other countries. *International Social Science Bulletin*, 1951, 3, 561-578.

Piaget, J. *Play, dreams and imitation in childhood*.

Translated by C. Gattegno & F. M. Hodgson. New York: Norton, 1962. Original French edition, 1945; First published in English, 1951.

London: Routledge & Kegan Paul, 1951.

London: Heinemann, 1951.

New York: Norton, 1952.

New York: Norton, 1962. (bp)

Piaget, J. The right to education in the modern world. In UNESCO. *Freedom and culture*. New York: Columbia University Press, 1951. Pp. 67-116.

Piaget, J. *The child's conception of number*.

Translated by C. Gattegno & F. M. Hodgson. New York: Norton, 1965. Original French edition, 1941; First published in English, 1952.

London: Routledge & Kegan Paul, 1952, 1961.

New York: Humanities Press, 1952.

New York: Humanities Press, 1962, 1964.

London: Routledge & Kegan Paul, 1965. (pb)

New York: Norton, 1965. (bp)

Piaget, J. Jean Piaget. In E. G. Boring *et al.* (Eds.), *A history of psychology in autobiography*. Vol. IV Barre: Clark University Press, 1952. Pp. 237-256.

Piaget, J. *The origin of intelligence in children*.

Translated by Margaret Cooke. New York: Norton, 1963. Original French edition, 1936; First published in English, 1952.

New York: International Universities Press, 1952, 1966.

London: Routledge & Kegan Paul, 1953.

New York: Norton, 1963. (pb).

Piaget, J. Genetic psychology and epistemology. *Diogenes*, 1953, 1, 49-63.

Piaget, J. How children form mathematical concepts. *Scientific American*, 1953, 189(20), 74-78.

Piaget, J. *Logic and psychology*.

Translated by W. Mays & F. Whitehead. (Based on lectures delivered at the University of Manchester, 1952.) New York: Basic Books, 1957. Original French edition, 1953; First published in English, 1953.

Manchester: The University Press, 1953.

New York: Basic Books, 1957.

- Piaget, J. *The construction of reality in the child*.
Translated by Margaret Cooke. New York: Ballantine, 1971. Original French edition, 1937; First published in English, 1954.
New York: Basic Books, 1954.
London: Routledge & Kegan Paul, 1955.
New York: Ballantine, 1971. (pb)
- Piaget, J. Language and thought from the genetic point of view. *Acta Psychologica*, 1954, 10, 51-60.
- Piaget, J. Problems of consciousness in child psychology: Developmental changes in awareness.
Translated by E. Meyer. In H. A. Abramson (Ed.), *Problems in consciousness*. Transactions of the Fourth Conference, 1953, Princeton, N. J. New York: J. Macey Foundation, 1954. Pp. 136-177.
- Piaget, J. The development of time concepts in the child. In R. H. Hoch, & J. Zubin (Eds.), *Psychopathology of childhood*. New York: Grune & Stratton, 1955. Pp. 34-44.
- Piaget, J. Perceptual and cognitive (or operational) structures in the development of the concept of space in the child. *Acta Psychologica*, 1955, 11, 41-46.
- Piaget, J., & Inhelder, B. *The child's conception of space*.
Translated by F. J. Langton & J. L. Lunzer. New York: Norton, 1967.
Original French edition, 1948; First published in English, 1956.
London: Routledge & Kegan Paul, 1956.
New York: Humanities Press, 1963.
New York: Norton, 1967. (pb)
- Piaget, J. Some impressions of a visit to Soviet psychologists. *American Psychologist*, 1956, 11, 343-345.
- Piaget, J. Some impressions of a visit to Soviet psychologists. *Acta Psychologica*, 1956, 12, 216-219.
- Piaget, J. The child and modern physics. *Scientific American*, 1957, 196(Mar.), 46-51.
- Inhelder, B. & Piaget, J. *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*.
Translated by Anne Parsons & S. Milgram. New York: Basic Books, 1958. Original French edition, 1955; First published in English, 1958.
- Piaget, J., Vinh-Bang, & Matalon, B. Note on the law of the temporal maximum of some optico-geometric illusions. *American Journal of Psychology*, 1958, 71, 277-282.

Piaget, J. Peering into the mind of a child. *The UNESCO Courier*, 1959, 12, 4-7.

Piaget, J., Inhelder, B., & Smeminska, A. *The child's conception of geometry*.

Translated by E. A. Lunzer. New York: Basic Books, 1960. Original French edition, 1948; First published in English, 1960.

London: Routledge & Kegan Paul, 1960.

New York: Basic Books, 1960.

Piaget, J. Individual and collective problems in the study of thinking. *Annals of the New York Academy of Sciences*, 1960, 91, 22-37.

Piaget, J. The genetic approach to the psychology of thought. *Journal of Educational Psychology*, 1961, 52, 275-281.

Piaget, J. The language and thought of the child. In *Classics in psychology*. New York: Thorno Shipley, 1961. Pp. 994-1031.

Piaget, J. *Comments on Vygotsky's critical remarks concerning "The language and thought of the child" and "Judgment and reasoning in the child"*. Cambridge, Mass.: M.I.T. Press, 1962.

Piaget, J. Three lectures. (The stages of the intellectual development of the child; The relation of affectivity to intelligence in the mental development of the child; Will and action.) *Bulletin of the Menninger Clinic*, 1962, 26, 120-145.

Piaget, J. The attainment of variance and reversible operations in developmental thinking. *Scientific American*, 1963, 30, 283-299.

Piaget, J. Cognitive development in children: The Piaget papers. (Development and learning; The development of mental imagery; Mother structures and the notion of number; Relations between the notions of time and speed in children.) In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered*. A Report of the Conference on Cognitive Studies and Curriculum Development, March 1964. Ithaca: Cornell University Press, 1964.

Piaget, J. Development and learning. *Journal of Research in Science Teaching*, 1964, 2, 176-186.

Inhelder, B., & Piaget, J. *The early growth of logic in the child: Classification and seriation*.

Translated by E. A. Lunzer & D. Papert. New York: Norton, 1969. Original French edition, 1959; First published in English, 1964.

London: Routledge & Kegan Paul, 1964.

New York: Harper & Row, 1964.

New York: Norton, 1969. (pb)

- Piaget, J. Psychology and philosophy.
Translated by R. Howard. In Wohlman & Nagel (Eds.), *Scientific psychology, Principles and approach*. New York: Basic Books, 1965. Pp. 28-43.
- Piaget, J. Foreward to M. Almy *et al.*, Young children's thinking: Studies of some aspects of Piaget's theory. New York: Teacher's College Press, Columbia University, 1966.
- Beth, E. W., & Piaget, J. *Mathematical epistemology and psychology*. New York: Gordon & Breach, 1966. Original French edition, 1961; First published in English, 1966.
- Piaget, J. Response to Brian Sutton-Smith. *Psychological Review*, 1966, 73, 111-112.
- Piaget, J. Time perception in children.
Translated by B. Montgomery. In J. P. Frazer (Ed.), *The voices of time*. New York: Braziller, 1966. Pp. 202-216.
- Piaget, J. Cognitions and conservations. *Contemporary Psychology*, 1967, 12, 532-533.
- Piaget, J. *Six psychology studies*. Edited by D. Elkin;
Translated by Anita Tenzer. New York: Vintage Books, 1969.
Original French edition, 1964; First published in English, 1967.
New York: Random House, 1967.
London: The University of London Press, 1969.
New York: Vintage Books, 1969. (pb)
- Piaget, J. *Studies in cognitive growth* (Bruner *et al.*):
Review. *Contemporary Psychology*, 1967, 12, 532-533.
- Piaget, J., *et al.* *Experimental psychology: Its scope and method*. Vol. I. *History and method*. New York: Basic Books, 1968.
- Piaget, J. *On the development of memory and identity*. (Heinz Werner Lecture Series, Clarke University, Worcester, Mass. Vol. 2.) New York: Columbia University Press, 1968.
- Piaget, J. Quantification, conservation and nativism. *Science*, 1968, 162, 976-979.
- Piaget, J. *The child's conception of movement and speed*.
Translated by G. E. T. Holloway & M. J. Mackenzie. New York: Ballantine, 1971. Original French edition, 1946; First published in English, 1969.
New York: Basic Books, 1969.
London: Routledge & Kegan Paul, 1970.
New York: Ballantine, 1971. (pb)

Piaget, J. *The child's conception of time.*

Translated by A. J. Pomerans. New York: Ballantine, 1971. Original French edition, 1927; First published in English, 1969.

London: Routledge & Kegan Paul, 1969.

New York: Basic Books, 1970.

New York: Ballantine, 1971. (pb)

Piaget, J. *Mechanisms of perception.*

Translated by G. N. Seagram. London: Routledge & Kegan Paul, 1969. Original French edition, 1961; First published in English, 1969.

London: Routledge & Kegan Paul, 1969.

New York: Basic Books, 1969.

Piaget, J., & Inhelder, B. The gaps in empiricism. In A. Koestler (Ed.), *Beyond reductionism: The Alpbach Symposium*, 1968. London: Hutchinson, 1969.

Piaget, J. & Inhelder, B. *The psychology of the child.*

Translated by Helen Weaver. New York: Basic Books, 1969. Original French edition, 1966; First published in English, 1969.

New York: Basic Books, 1969.

New York: Harper Torchbooks, 1969. (pb)

London: Routledge & Kegan Paul, 1969.

Piaget, J. *Genetic epistemology.*

Translated by Eleanor Duckworth. Woodbridge Lectures, 1968. New York: Norton, 1971. First published in English, 1970.

New York: Columbia University Press, 1970.

New York: Norton, 1971. (pb)

Piaget, J. *Genetic psychology.*

Translated by Eleanor Duckworth. Ithaca: Cornell University Press, 1970.

Piaget, J. Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's Manual of child psychology*, Vol. I. New York: Wiley, 1970. Pp. 703-732.

Piaget, J. *Science of education and the psychology of the child.* New York: Viking, 1971. First published in English, 1970.

New York: Orion Press, 1970.

New York: Viking, 1971 (pb)

Piaget, J. *Structuralism.*

Translated by Channah Maschler. New York: Harper Torchbooks, 1971. Original French edition, 1968; First published in English, 1970.

New York: Basic Books, 1970.

New York: Harper Torchbooks, 1971. (pb)

Piaget, J. *Biology and knowledge: An essay on the relations between*

- organic regulations and cognitive processes*. Edinburgh: Edinburgh University Press, 1972. First published in English, 1971.
Chicago: University of Chicago Press, 1971.
Edinburgh: Edinburgh University Press, 1972.
- Piaget, J. *et al. Mental imagery in the child: A study of the development of imaginal representation*.
Translated by P. A. Chilton. London: Routledge & Kegan Paul, 1971.
Original French edition, 1966; First published in English, 1971.
- Piaget, J. The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer (Eds.), *Measurement and Piaget*, Proceedings of the CTB/McGraw Hill Conference on Ordinal Scales of Cognitive Development. New York: McGraw-Hall, 1971. Pp. 1-11.
- Piaget, J. Comments on mathematical education. Paper presented at the International Conference for Mathematics Education, Exeter, 1972.
- Piaget, J. *Insights and illusions of philosophy*.
Translated by W. Mays. London: Routledge & Kegan Paul, 1972.
First published in English, 1972.
- Piaget, J. Intellectual evolution from adolescence to adulthood. *Human Development*, 1972, 15, 1-12.
- Piaget, J. Play and development. In M. W. Piers (Ed.), *Play and development: A symposium*. New York: Norton, 1972.
- Piaget, J. *The principles of genetic epistemology*.
Translated by Mays Wolfe. New York: Basic Books, 1972. First published in English, 1972.
- Piaget, J. Problems of equilibration. In C. F. Nodine, J. M. Gallagher, & R. H. Humphreys (Eds.), *Piaget and Inhelder: On equilibration*. Philadelphia: The Jean Piaget Society, 1972. Pp. 1-20.
- Piaget, J. *Psychology and epistemology: Towards a theory of knowledge*.
Translated by Arnold Rosin. New York: Viking, 1972. First published in English, 1972.
- Piaget, J. The role of imitation in the development of representational thought. *International Journal of Mental Health*, 1972-3, 1, 67-74.
- Piaget, J. *The child and reality: Problems of genetic psychology*.
Translated by Arnold Rosin. New York: Grossman, 1973. Original French edition, 1972; First published in English, 1973.
- Piaget, J., & Inhelder, B. *Memory and intelligence*.
Translated by Arnold J. Pomerans. London: Routledge & Kegan Paul, 1973. Original French edition, 1968; First published in English, 1973.